第II部分　方　　法

非线性规划的实际兴趣，集中在有效地找出给定问题的数值解. 单纯形方法在线性规划中起着普遍适用的作用，与这种情况相反，对于获得非线性规划的最优解来说，还没有这种统一的处理. 这样，单纯形方法能够有效地解含有几千个变量和几百个约束的线性规划问题，而如何求解少数变量的无约束非线性函数的极小化问题，却仍是许多学者关切的一个重要问题. 从算法观点来研究无约束最优化，将是第II部分前四章的主题.

我们从单变量实函数极小化的最简单情形开始. 然后，我们提出不用导数的方法，也就是说，提出一些寻找极值的算法，而不利用需找极值的那个函数的导数. 这些方法是重要的，因为在大多数实际应用中，导数或者是不能使用的，或者是难于计算的. 不理想的是，对于这种问题所设计的现有方法不是那么有效，所以在这个领域还有许多工作要做. 另一方面，1960 年以来，在发展无约束最优化问题的带有导数的有效算法方面，出现了重大进展. 在利用新一代计算机的某些平行计算特征而设计数值方法方面，其最初步骤也将在这里提到.

无约束最优化方法的研究是重要的，因为求解约束非线性规划问题的最成功的一类算法是惩罚方法，它基于将约束问题转化为无约束问题. 另一些方法直接地处理约束，它考虑到约束，以修正从变量空间中的一个点到另一个点的移动. 最后三章研究约束最优化，以结束本书的第二部分.

第 8 章

一维最优化

这一章集中研究的问题是：对实直线 R 或 R 的某一子集上的单变量函数 f，寻找它的极小值．在这一章中单独地讨论一维最优化，这是有几个理由的．首先，在 R^n 上无约束最优化的某些理论背景和数值背景能方便地用一维情形来说明；其次，对单变量函数的一些有趣的搜索方法显然不能完整地推广到在 R^n 上寻求函数极值的问题；最后，n 维问题的某些迭代方法包含了在 R^n 中沿某些方向寻找极值的步骤，而这些步骤实际上等价于一维最优化．

因此，我们研究单个实变量 x 的实函数 f 的极小化问题．在第二章中我们已经看到，如果 f 在 x^* 取得局部极值且如果在 x^* 可以微分，那么 f 的一阶导数在 x^* 必须等于零，即局部极值点是下列方程的解：

$$f'(x)=0. \tag{8.1}$$

我们可以尝试去求解这个一般的非线性方程，得到它的所有解，如果局部极小值点存在，则必在这些解中间．如果 f 的二阶导数亦可用，那么在(8.1)的解处计算 f'' 的值，我们通常就能决定哪一些解对应于极小．

我们将讨论沿一条直线（即一维）寻找实值函数极小值点的三种方法．首先是用于求解(8.1)的 Newton 法和某些有关方法．第二种方法是用低阶（通常不超过三阶）多项式逼近寻求极小的函数．于是，多项式的极小可用分析的方法求得，同时又能得到一个新的近似多项式．继续这个方法，我们能得到极小值点真实位置的近似值．单变量函数的第三种优选法属于“直接法”．这里，在它的定义域中不同点的函数值按照某种系统的办法进行计算，同时，根据这些函数的取值和某些适当的假设，我们可以使极小值点

的位置划入某一个小区域内.

8.1 Newton 法

在多数情况下，无法求得方程(8.1)的解析解，所以这个方程必须采用某个迭代法求解，这个方法产生点列 $\{x^k\}$ 和导数值序列 $\{f'(x^k)\}$，使得导数值序列的极限为零. 解非线性方程的经典方法是 Newton 法，在这里这个方法要求知道二阶导数 $f''(x)$ 的值. Newton 法的思想是，设 x^k 是(8.1)的解的猜测点，在 x^k 的附近使 f' 线性化，并且求得使这个线性函数为零的点. 这个点就是下一个猜测点 x^{k+1}. 形式上，令 x^k 为 $f'(x)=0$ 的解的当前猜测点. 线性方程

$$f'(x^k)+f''(x^k)(x-x^k)=0 \tag{8.2}$$

的根为

$$x^{k+1}=x^k-\frac{f'(x^k)}{f''(x^k)}, \tag{8.3}$$

这个迭代方法也在图 8.1 中阐明.

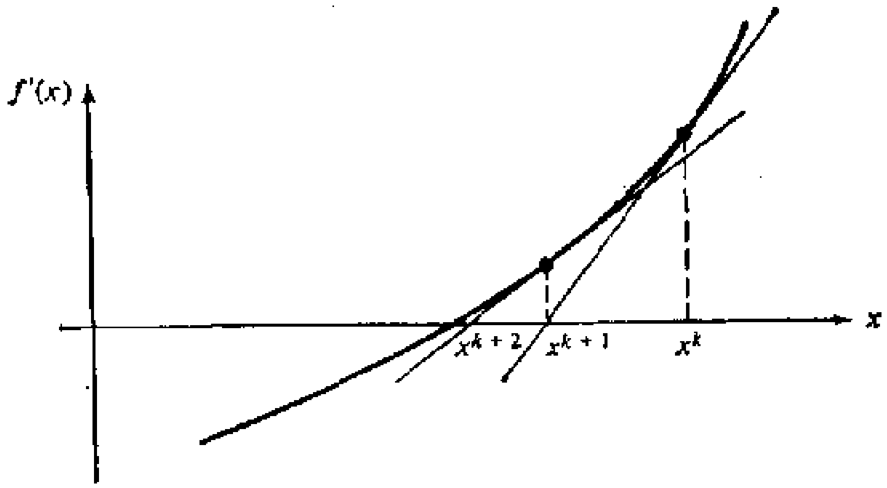


图 8.1 Newton 法

我们现在证明，在适当的假设下，由上述公式产生的点列 $\{x^k\}$ 收敛于 (8.1) 的一个解. 下面关于 Newton 法的讨论完全仿照 Goldstein 的工作[7]，在他的著作中读者能够看到更完全的分析.

设 $S=[a, b]$ 是实直线上的一个闭区间，ϕ 是 S 上的连续实

函数. 又设 T 是 ϕ 的值域. 如果 $T\subset S$, 那么 ϕ 称为从 S 到自身的映射. 如果点 $x\in S$, 使得 $\phi(x)=x$, 则点 x 称为 ϕ 的**不动点**. 如果存在一个正数 $q<1$, 使得对于 S 中任何两点 x^1 与 x^2, 成立

$$|\phi(x^1)-\phi(x^2)|\leqslant q|x^1-x^2|, \tag{8.4}$$

那么从 S 到自身的映射 ϕ 称为**压缩映射**. 这样, 我们有如下预备性结果.

引理 8.1

假设 ϕ 在 $S=[a,\ b]\subset R$ 上是压缩映射, 且 $x^0\in S$, $x^{k+1}=\phi(x^k)$, 那么, 存在 ϕ 的唯一不动点 $\bar{x}$, 序列 $\{x^k\}$ 收敛于 $\bar{x}$, 并且满足

$$|x^{k+1}-\bar{x}|\leqslant(q)^{k+1}|x^0-\bar{x}|, \quad k=0,\ 1,\ \cdots. \tag{8.5}$$

【证明】 将 S 映射到自身的每一个连续函数 ϕ 具有一个不动点[7]. 现在有

$$|x^1-\bar{x}|=|\phi(x^0)-\phi(\bar{x})|\leqslant q|x^0-\bar{x}|. \tag{8.6}$$

因此, 当 $k=0$ 时 (8.5) 成立. 假设当 $k=K\geqslant 1$ 时 (8.5) 成立. 那么

$$|x^{K+1}-\bar{x}|\leqslant(q)^{K+1}|x^0-\bar{x}|, \tag{8.7}$$

由 (8.6) 和 (8.7), 得

$$\begin{aligned}|x^{K+2}-\bar{x}|&=|\phi(x^{K+1})-\phi(\bar{x})|\leqslant q|x^{K+1}-\bar{x}|\\&\leqslant(q)^{K+2}|x^0-\bar{x}|.\end{aligned} \tag{8.8}$$

这样 (8.5) 对任意 $k=0,\ 1,\ \cdots$ 成立. 因为 $q<1$, 所以公式 (8.5) 蕴涵序列 $\{x^k\}$ 收敛于 $\bar{x}$. 最后, 必须指出 $\bar{x}$ 是唯一的. 假设有两个不同的不动点 $\bar{x}^1$ 和 $\bar{x}^2$. 那么

$$0<|\bar{x}^1-\bar{x}^2|=|\phi(\bar{x}^1)-\phi(\bar{x}^2)|\leqslant q|\bar{x}^1-\bar{x}^2|, \tag{8.9}$$

因为 $q<1$, 上式是一个矛盾. 】

下面的引理给出 ϕ 是压缩映射的充分条件.

引理 8.2

假设 ϕ 将 $S=[a,\ b]\subset R$ 映射到自身, 并设 ϕ 在 S 上具有连续导数, 如果对于每一个 $x\in S$ 有 $|\phi'(x)|<1$, 那么 ϕ 是压缩映射.

【证明】 设 x^1 与 x^2 是 S 中的任意两点, 那么由中值定理[4],

$$\phi(x^1)=\phi(x^2)+\phi'(\tilde{x})(x^1-x^2), \tag{8.10}$$

其中 $\tilde{x}$ 是 x^1 与 x^2 之间的某一点. 因此,

$$|\phi(x^1)-\phi(x^2)|=|\phi'(\tilde{x})||x^1-x^2|, \tag{8.11}$$

同时由 ϕ' 的假设, 我们取 q 等于 $|\phi'(x)|$ 在 S 上的最大值, 就知道 ϕ 是压缩的. ▌

借助于这两个引理, 我们能证明关于某些求根方法收敛性的一个一般结果.

定理 8.3

设 h 和 γ 是 $S=[a, b]\subset R$ 上的连续可微函数. 假设

$$h(a)h(b)<0, \tag{8.12}$$

同时对于一切 $x\in S$, 有

$$\gamma(x)>0, \quad h'(x)>0. \tag{8.13}$$

$$0\leqslant 1-[\gamma(x)h(x)]'\leqslant q<1. \tag{8.14}$$

令 $x^0\in S$, 且

$$x^{k+1}=x^k-\gamma(x^k)h(x^k), \quad k=0, 1, \cdots, \tag{8.15}$$

那么序列 $\{x^k\}$ 收敛于 $h(x)=0$ 的解 $\bar{x}$.

【证明】 令

$$\phi(x)=x-\gamma(x)h(x), \tag{8.16}$$

$$\phi'(x)=1-[\gamma(x)h(x)]', \tag{8.17}$$

从(8.14)可知, 对于一切 $x\in S$, 有

$$0\leqslant\phi'(x)\leqslant q<1, \tag{8.18}$$

并且 ϕ 在 S 上是单调非减的. 由(8.12)和(8.13)可知, 在 S 上 h 是单调增加的, 并且 $h(a)<0$, $h(b)>0$. 因此 $\phi(a)>a$, $\phi(b)<b$. 由 ϕ 的单调性推知, 对于一切 $x\in S$ 有 $a<\phi(x)<b$. 因此 ϕ 将 S 映射到自身. 此外 $|\phi'(x)|<1$, 由引理 8.2 知函数 ϕ 在 S 上是压缩的. 从(8.15)和(8.16)有 $x^{k+1}=\phi(x^k)$, $k=0, 1, \cdots$. 利用引理 8.1, ϕ 有唯一的不动点 $\bar{x}\in S$. 最后, 注意到当且仅当 $h(\bar{x})=0$ 时 $\bar{x}$ 才是 ϕ 的不动点. 因此, $\{x^k\}$ 收敛于 $h(x)=0$ 的解. ▌

现在我们能够阐明上述 Newton 法收敛的充分条件.

推论 8.4

设 $h(x)=f'(x)$, $\gamma(x)=1/f''(x)$, 其中 f 是 $S=[a, b]$ 上的三

次连续可微函数．假定在定理 8.3 中给出的关于 h 和 γ 的假设是成立的．那么(8.15)就是 Newton 法的迭代公式，并且$\{x^k\}$收敛于 f 的不动点．

【证明】 直接从定理 8.3 推得． 】

注意，上述关于 γ 的条件蕴涵着 f 在 S 上是严格凸的（相当强的假设）．虽然条件包含 f 的三阶导数，但在算法中是不用的．

除了证明某一算法的收敛性外，知道**收敛速率**也是重要的．假定存在一序列 $\{x^k\}$，$x^k\in R^n$，收敛于点 $\bar{x}$，并且对于所有充分大的 k 有 $x^k\neq\bar{x}$．如果存在一个数 p 和 $\alpha\neq 0$，使得

$$\lim_{k\to\infty}\frac{\|x^{k+1}-\bar{x}\|}{\|x^k-\bar{x}\|^p}=\alpha, \tag{8.19}$$

那么 p 称为序列 $\{x^k\}$ 的收敛阶，$\|x^k-\bar{x}\|$称为第 k 次逼近的误差．如果 $p=1$，序列 $\{x^k\}$ 的收敛速率称为是**线性的**．如果 $p>1$，那么收敛速率称为是**超线性的**．对于 $p=2$ 的特殊情况，收敛速率称为是**二阶的**．线性收敛的更一般的概念如下：存在一个自然数 K，以及数 β 和 q，使得 $0\leqslant q<1$，并且满足

$$\|x^k-\bar{x}\|\leqslant\beta(q)^k,\quad k=K,\ K+1,\ \cdots. \tag{8.20}$$

(8.19)给出的超线性收敛性的较弱的定义为

$$\lim_{k\to\infty}\frac{\|x^{k+1}-\bar{x}\|}{\|x^k-\bar{x}\|}=0. \tag{8.21}$$

于是我们有下面的定理．

定理 8.5

假设定理 8.3 和推论 8.4 的假设成立，由 Newton 法产生的序列 $\{x^k\}$，$x^k\in R$，收敛于点 $\bar{x}$，$\bar{x}$ 满足 $h(x)=0$，那么收敛的速率是二阶的．

【证明】 当且仅当 $\bar{x}$ 是

$$\phi(x^k)=x^k-\frac{h(x^k)}{h'(x^k)} \tag{8.22}$$

的不动点时，点 $\bar{x}$ 满足 $h(x)=0$．由中值定理，

$$x^{k+1}-\bar{x}=\phi(x^k)-\phi(\bar{x})=\phi'(\xi^k)(x^k-\bar{x}), \tag{8.23}$$

其中 ξ^k 位于 x^k 与 $\bar{x}$ 之间．于是

$$|x^{k+1}-\bar{x}|=\frac{|h''(\xi^k)h(\xi^k)|}{[h'(\xi^k)]^2}|x^k-\bar{x}|, \tag{8.24}$$

同时，

$$\begin{aligned}|h(\xi^k)|=|h(\xi^k)-h(\bar{x})|&=|h'(\eta^k)||\xi^k-\bar{x}|\\ &\leqslant|h'(\eta^k)||x^k-\bar{x}|,\end{aligned} \tag{8.25}$$

其中 η^k 位于 ξ^k 与 $\bar{x}$ 之间．因此

$$|x^{k+1}-\bar{x}|\leqslant\frac{|h''(\xi^k)h'(\eta^k)|}{[h'(\xi^k)]^2}|x^k-\bar{x}|^2. \tag{8.26}$$

设

$$\beta=\sup_k\frac{|h''(\xi^k)h'(\eta^k)|}{[h'(\xi^k)]^2}, \tag{8.27}$$

那么

$$|x^{k+1}-\bar{x}|\leqslant\beta|x^k-\bar{x}|^2, \tag{8.28}$$

即(8.19)成立． ▌

用差商来近似二阶导数 $f''(x)$

$$f''(x^k)\cong\frac{f'(x^k)-f'(x^{k-1})}{x^k-x^{k-1}}, \tag{8.29}$$

就能得到密切相关的一个求根方法．将上式代入(8.3)，得到割线法的迭代公式

$$x^{k+1}=x^k-\frac{f'(x^k)(x^k-x^{k-1})}{f'(x^k)-f'(x^{k-1})}. \tag{8.30}$$

能够证明[10]，对于割线法有

$$\lim_{k\to\infty}\frac{|x^{k+1}-\bar{x}|}{|x^k-\bar{x}|^\tau}=\left|\frac{2f''(\bar{x})}{f'''(\bar{x})}\right|^{1/\tau}, \tag{8.31}$$

其中假设 f''' 不为零，同时 $\tau=(1+\sqrt{5})/2\cong1.618$ 是方程 $t^2=t+1$ 的解．因此，对于充分大的 k，割线法是近似于1.618阶的超线性方法．这里有趣的是要注意，联系着另一个一维最优化方法，数 τ 在本章以后部分中将再次出现．

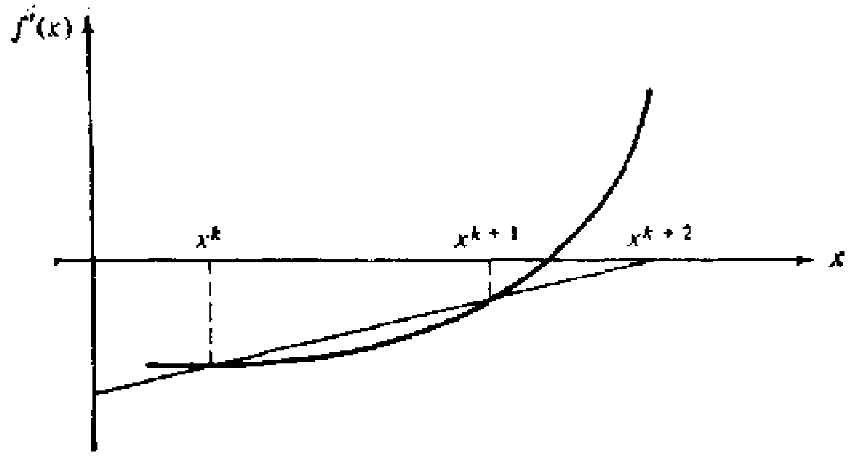


图 8.2 割线法

也许提醒下列一点是合适的：通过前面定义的收敛速率来比较两种算法，是可能会引起误解的．例如，割线法的渐近收敛速率 τ 是关于函数赋值一次的，因为在每次迭代中需要求根的函数只计算一次．另一方面，Newton 法的二阶收敛速率是关于函数赋值两次的，因为在每次迭代中函数及其一阶导数均必须赋值．

8.2 多项式逼近法

下面将研究两种最广泛采用的多项式逼近方法，即二次方法和三次方法．二次方法的基础是用二次函数 ϕ 逼近要寻求极小的函数 f，函数 ϕ 由下式给出：

$$\phi(x)=a+bx+c(x)^2, \tag{8.32}$$

这时假定在三个点 x^1、x^2 和 x^3 我们算得 f 的值，并且 $x^1<x^2<x^3$．

现令 $\phi(x^i)=f(x^i)$，而 $i=1, 2, 3$，我们能解得系数 a，b 和 c．二次函数 ϕ 的极小值点（如果它存在极小的话）能解析地从 $\phi'(x)=0$ 求得，并且，作为 f 的极小值点的第一次逼近，我们得到点 $\hat{x}$ 如下：

$$\hat{x}=-\frac{b}{2c}, \tag{8.33}$$

此处假定 $c>0$．如果 $c<0$，二次函数实际上是具有极大值的抛物线，因而求得的点 $\hat{x}$ 是不能用的．能保证 c 是正的一种情况是

$$f(x^1)>f(x^2) \quad 和 \quad f(x^2)<f(x^3). \tag{8.34}$$

如果上式成立，我们就能确信，f 的一个局部极小值点已划入 x^1 与 x^3 之间．因而这样的划界是非常合适的，我们很快将介绍这样做的步骤，但是先继续讨论二次方法，这里假定(8.34)是满足的．这样求得的 ϕ 的极小值点也满足

$$f(x^1)>\phi(\hat{x}) \quad 和 \quad \phi(\hat{x})<f(x^3). \tag{8.35}$$

同时，我们能够找出新的二次逼近和近似极小值点，办法是选择三点的新集合如下：计算值 $f(\hat{x})$，并从计算 f 值的四个点中选取一个使 f 值取最小的点作为新的 x^2（新 x^2 或者是 $\hat{x}$ 或者是"旧"的 x^2）．令新的 x^1 和 x^3 各是新的 x^2 近旁左右两个点，重复迭代．当在预测的极小值点的真实函数值和逼近函数值之差小于某一容许的 $\varepsilon>0$ 时，也就是如果

$$|f(\hat{x})-\phi(\hat{x})|<\varepsilon, \tag{8.36}$$

或者当两个或更多个相继迭代的极小值点之间距离小于某一预先给定的距离时，这个算法就终结．然而，对于上述二次逼近法，即使不等式(8.34)满足，至少有一种情况它不能收敛于 f 的极小．如果 $\hat{x}=x^2$，即使 x^2 可能不是 f 的局部极小值点，算法也不产生新点．在这种退化情况下，为了继续进行计算，需要使 $\hat{x}$ 作某种扰动．

不等式(8.34)对于预测 ϕ 的极小是充分的(除了在退化情况外)．但如(8.34)不成立，也能预测这样一个极小．例如，如果 $f(x^1)>f(x^2)>f(x^3)$，并且 $f(x^2)$ 位于联结 $f(x^1)$ 和 $f(x^3)$ 的直线段的下方，相应的二次逼近仍然可能在 $\hat{x}>x^3$ 处具有极小．然而，这个要预测的极小可能离开第一次三个点的右边很远，以致用 ϕ 来近似 f 可能是很差的．如果预测的极小出现这种外推的情形，只要引入极大步长，便可以克服这个困难．另一方面，点 x^1, x^2, x^3 是可能离 f 的极小很远的，因此这种用外推求 ϕ 的极小值点能加速迭代过程．

要求出三点使得(8.34)得到满足的划界方法叙述如下．选择任一初始点 x^0 并计算 $f(x^0)$．其次，在点 x^0+h^0 算出 f 的数值，其中 $h^0>0$ 是某一预先给定的步长．如果

$$f(x^0)>f(x^0+h^0), \tag{8.37}$$

令 $x^1=x^0+h^0$，又令 $h^1=\alpha h^0$ 为新的步长，其中 $\alpha>1$．现在计算 $f(x^1+h^1)$．如果

$$f(x^1)>f(x^1+h^1), \tag{8.38}$$

令 $x^2=x^1+h^1$ 和 $h^2=\alpha h^1$．继续这个办法，直到 f 在最新达到的点 x^k+h^k 上的值高于 $f(x^k)$．

如果(8.37)不成立，令 $x^1=x^0$ 和 $h^1=-\beta h^0$，其中 $0<\beta<1$，算出数值 $f(x^1+h^1)$．如果

$$f(x^1+h^1)>f(x^1), \tag{8.39}$$

我们就结束，因为要划的区间 $[x^0-\beta h^0, x^0+h^0]$ 已经得到．否则，令 $x^2=x^1+h^1$，$h^2=\alpha h^1$ 并计算 $f(x^2+h^2)$，如此继续下去，直到函数出现增加．不论怎样，最后得到的三点将满足(8.34)，于是二次方法就能开始．α 和 β 的值宁可选得使 $\alpha\beta<1$．应该指出，划界过程本身就能作为确定极小值点的方法．

在图 8.3 中阐明了上述情况，每当获得函数增加，就颠倒搜索方向（小箭头由第 k 次函数值指向下一次函数值）．在图中采用 $\alpha=2$，$\beta=\frac{1}{4}$，同时已做了 11 个函数值计算．关于这个方法的细节，读者可参考[12]．注意，如果函数 f 的一阶导数已知，则二次方法和划界方法均能修正和改善．

下面研究的多项式逼近方法是**三次方法**，这方法采用三阶多

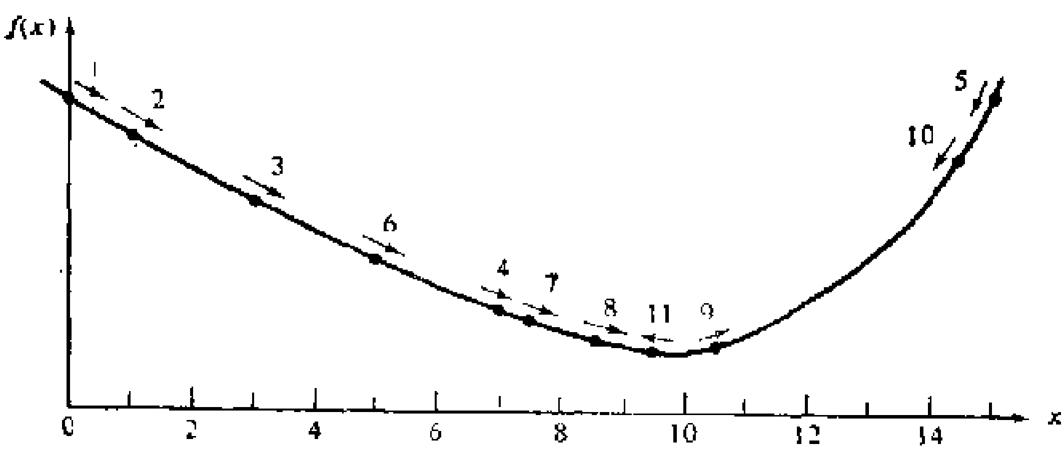


图 8.3　用划界法求函数极小

项式

$$\phi(x)=a+bx+c(x)^2+d(x)^3 \tag{8.40}$$

来逼近寻求极小的给定函数 f. 这里介绍的方法原来由 Davidon[6] 导出. 算法基于假设 f 的一阶导数是可用的, 但这种方法也能容易地修改使之不必计算导数. 从任意点 $x^1\in R$ 开始, 计算 $f(x^1)$, $f'(x^1)$. 假定 $f'(x^1)$是负的, 那么我们通过某一个迭代过程找一点 x^2, $x^2>x^1$, 使得 $f'(x^2)$是非负的或者 $f(x^2)>f(x^1)$. 系数 a, b, c 和 d 能从求解下列四个变量四个线性方程的方程组而得到:

$$\left.\begin{aligned}f(x^1)&=a+bx^1+c(x^1)^2+d(x^1)^3,\\ f'(x^1)&=b+2cx^1+3d(x^1)^2,\\ f(x^2)&=a+bx^2+c(x^2)^2+d(x^2)^3,\\ f'(x^2)&=b+2cx^2+3d(x^2)^2.\end{aligned}\right\} \tag{8.41}$$

Davidon 发现这些方程的解能够通过简单的变量代换迅速求出. 用下列方程定义一个新的变量 z 和新的函数 g 和 ψ:

$$z=x-x^1, \tag{8.42}$$

$$g(z)=f(x^1+z),\quad \psi(z)=\phi(x^1+z), \tag{8.43}$$

那么三次函数 ψ 关于 z 的一阶导数为

$$\psi'(z)=g'(0)-\frac{2z}{\lambda}(g'(0)+\alpha)+\frac{(z)^2}{(\lambda)^2}(g'(0)+g'(\lambda)+2\alpha), \tag{8.44}$$

其中 $\lambda=x^2-x^1$, 且

$$\alpha=\frac{3[g(0)-g(\lambda)]}{\lambda}+g'(0)+g'(\lambda), \tag{8.45}$$

同时使 $\psi'(z)$为零, 也就是预测为极小的点 $\hat{z}$ 由下式给出:

$$\hat{z}=\lambda(1-\beta), \tag{8.46}$$

其中

$$\beta=\frac{g'(\lambda)+[(\alpha)^2-g'(0)g'(\lambda)]^{1/2}-\alpha}{g'(\lambda)-g'(0)+2[(\alpha)^2-g'(0)g'(\lambda)]^{1/2}}. \tag{8.47}$$

设 ε 是某一个预先给定的容许值, 如果 $|g'(\hat{z})|<\varepsilon$, 那么过程结束; 否则算法必须采用新的两个点重复进行, 两点的选取方法类似于已介绍的二次逼近方法. 三次逼近方法中每次迭代的收敛速率

通常比二次方法要快．然而，三次方法中每次迭代所需的计算量比较大．

对于寻找有适当“好性质”的函数的极小值点——即这种函数能用二次或三次函数适当接近地近似(在某种意义下)——多项式逼近方法似乎是十分有效的．这些算法的效率在很大程度上依赖于要寻找极小的具体函数．下面介绍的方法寻找实直线上的小区间，在此区间上包含单峰函数(定义见下面)的极小值点．有趣的是，不管求极小的是哪一种具体的一维函数，只要属于单峰函数类，这些方法的效率将证明都是相同的．

8.3 直接法——Fibonacci 法和黄金分割法

我们从定义单峰函数开始，它对建立下面介绍的方法的收敛性和最优性都是非常重要的．一个实值函数 f 称为在闭区间 $L\subset R$ 上是单峰的，如果存在一个 $x^*\in L$，使得 x^* 是 f 在 L 上的极小值点，而且对于任意两点 $x^1\in L$, $x^2\in L$, $x^1<x^2$，我们有

$$\left.\begin{array}{l} x^2\leqslant x^* \text{ 蕴涵 } f(x^1)>f(x^2), \\ x^*\leqslant x^1 \text{ 蕴涵 } f(x^2)>f(x^1). \end{array}\right\} \tag{8.48}$$

注意，单峰函数不一定是连续的或可微的．严格凸函数及其许多推广都是单峰函数．单峰函数的一些例子如图 8.4 所示．

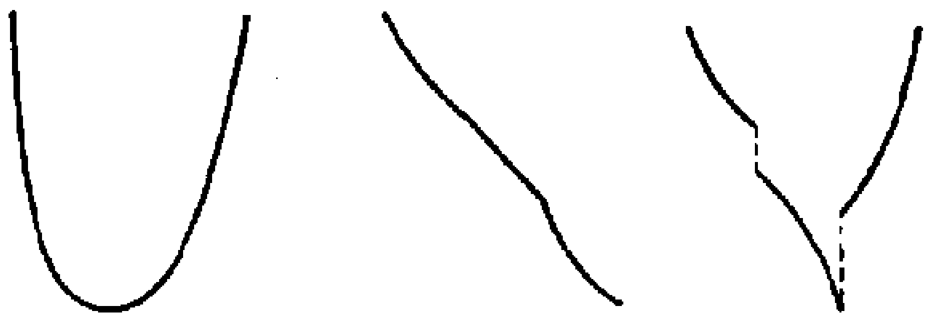


图 8.4 单峰函数

在下述方法中将充分利用单峰函数的一个非常有用的性质：通过在 L 内相异两点计算函数值，就可以划定极小的位置．分别以 l_1 和 r_1 表示区间 L 的左端点和右端点；即 $L=[l_1, r_1]$，或者

$$L=\{x:x\in R,\ l_1\leqslant x\leqslant r_1\}. \tag{8.49}$$

假定我们在 L 内有两点 x^1 和 x^2, $x^1<x^2$, 计算函数值得到 $f(x^1)<f(x^2)$. 从单峰的定义(8.48)推知 $x^*\in[l_1,\ x^2]$. 类似地, 如果 $f(x^1)>f(x^2)$, 那么必有 $x^*\in[x^1,\ r_1]$. 如果在 x^1 和 x^2 的函数值恰巧相等, 那么 $x^*\in[x^1,\ x^2]$, 但是, 为了简化起见, 我们认为 x^* 属于前述两个较大的区间中的一个. 不管怎样, 在计算头两个函数值之后, 位于 x^2 的右边或 x^1 的左边的 L 的一部分在以后搜索中可以划去. 如果留下的区间是 $[l_1,\ x^2]$, 那么它将包含点 x^1, 反之, 如果 $[x^1,\ r_1]$ 留下, 则它包含 x^2. 重记 L 留下的子区间为 $[l_2,\ r_2]$, 在 $[l_2,\ r_2]$ 上取一点异于已经出现的点. 在所取点上计算函数值 f. 将这一点上 f 的函数值与 $[l_2,\ r_2]$ 中另一点的函数值相比较, 我们能进一步缩小这个区间, 每一个这样的区间缩小称为一次迭代.

下面介绍的方法称为 **Fibonacci 法**, 因为它采用下式定义的 Fibonacci 数 F_k:

$$F_0=0,\ F_1=1,\ F_k=F_{k-1}+F_{k-2},\quad k=2,\ 3,\ \cdots. \tag{8.50}$$

因此, 前几个 Fibonacci 数为 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, ….

Fibonacci 法定义如下: 设 N 为函数 f 计值点的总数 (注意, 对于 N 个函数值, 具有 $N-1$ 次区间缩小或迭代). 在第 k 次迭代, 包含 x^* 的区间为 $[l_k,\ r_k]$ 对于 $k=1,\ 2,\ \cdots,\ N-1$, 比较下列两点的函数值:

$$x_k^1=l_k+\frac{F_{N-k}}{F_{N+2-k}}(r_k-l_k), \tag{8.51}$$

$$x_k^2=l_k+\frac{F_{N+1-k}}{F_{N+2-k}}(r_k-l_k), \tag{8.52}$$

这里除了 $k=1$ 以外, 总有一点 x_k^1 或 x_k^2 在前一次迭代中已经计算过函数值了.

注意, 点 x_k^1 和 x_k^2 对称地安置在区间 $[l_k,\ r_k]$ 中. 从(8.51)和(8.52)以及 Fibonacci 数的定义, 我们有

$$x_k^2-l_k=\frac{F_{N+1-k}}{F_{N+2-k}}(r_k-l_k)=\frac{F_{N+2-k}-F_{N-k}}{F_{N+2-k}}(r_k-l_k) \quad (8.53)$$

$$=r_k-l_k-\frac{F_{N-k}}{F_{N+2-k}}(r_k-l_k)=r_k-x_k^1, \quad (8.54)$$

(也可见图 8.5).

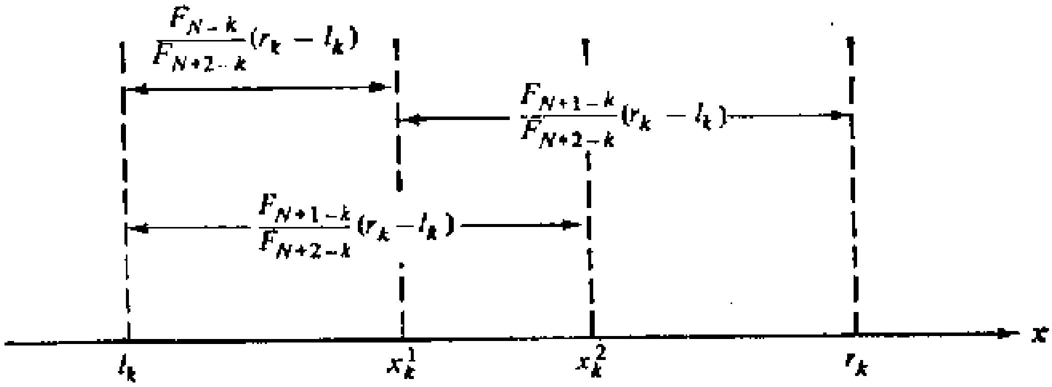


图 **8.5**　Fibonacci 法在对称点计算函数值

在最后一次迭代($k=N-1$)中，公式(8.51)和(8.52)中两个点 x_{N-1}^1 和 x_{N-1}^2 应当安置在

$$x_{N-1}^1=x_{N-1}^2=l_{N-1}+\frac{1}{2}(r_{N-1}-l_{N-1}), \quad (8.55)$$

从而不可能进一步缩减区间. 因此，为了通过最后的函数值计算使$[l_{N-1}, r_{N-1}]$缩小到它的长度的一半，我们把最后的点 x_{N-1}^1 和 x_{N-1}^2 放在离区间$[l_{N-1}, r_{N-1}]$中现有点的距离为 ε 的地方，这里 ε 取为要计算函数值的两相邻点之间的最小的实际距离，使得两个函数值是可区别的. 更准确地说，如果 $x_k^2-x_k^1\geqslant\varepsilon$，那么仅当 x^* 位于 x_k^1 与 x_k^2 之间时 $f(x_k^1)=f(x_k^2)$.

在计算了 N 次函数值后，包含 x^* 的区间的长度由下式给出:

$$r_N-l_N=\frac{r_1-l_1}{F_{N+1}}+\delta, \quad (8.56)$$

其中 δ 是零或者 ε. 因此我们能够通过 11 次函数值计算($F_{12}=144$)，在初始区间的 1% 以内划出任何单峰函数的极小值点，而通过 16 次函数值计算能在初始区间的 0.1% 以内划出.

例 8.3.1

试寻求 $f(x)=(x-3)^2$ 的极小值点，用 $N=4$ 次函数值计算,

并且假定 $L=[0, 10]$.

用 Fibonacci 法在搜索结束时，应该得到包含 f 的极小值点 $x^*=3$ 的最终区间，它的长度为

$$r_4-l_4=\frac{10-0}{5}+\varepsilon=2+\varepsilon. \tag{8.57}$$

搜索开始如下：

$$x_1^1=0+\frac{2}{5}(10-0)=4, \tag{8.58}$$

$$x_1^2=0+\frac{3}{5}(10-0)=6, \tag{8.59}$$

$$f(x_1^1)=1, \quad f(x_1^2)=9. \tag{8.60}$$

现在我们缩小初始区间，按前面确定的规则，得到 $l_2=0$, $r_2=6$. 进行第二次迭代，从(8.51)和(8.52)得

$$x_2^1=0+\frac{1}{3}(6-0)=2, \tag{8.61}$$

$$x_2^2=0+\frac{2}{3}(6-0)=4. \tag{8.62}$$

注意，在这情况下，点 x_2^2 重合于 x_1^1, 而在点 x_1^1 已经计算过 f 的值. 这里

$$f(x_2^1)=1, \quad f(x_2^2)=1. \tag{8.63}$$

虽然(8.63)指出 f 的极小值点位于 $x=2$ 和 $x=4$ 之间，我们有意不管这种意外情况(实际中很少出现)，而令 $l_3=2$, $r_3=6$. 于是

$$x_3^1=2+\frac{1}{2}(6-2)=4, \tag{8.64}$$

$$x_3^2=2+\frac{1}{2}(6-2)=4. \tag{8.65}$$

因为在 $x=4$ 的函数值已经计算过，不妨选取 $\varepsilon=0.01$, 得 $f(4.01)=1.0201$. 因此，如预测那样，包含 f 的极小值点的最终区间为 $[2, 4.01]$. 搜索过程也表明在图 8.6 中. ∎

Fibonacci 搜索法是由 Kiefer[9] 首先提出的，他也曾指出，在某种意义下这个方法是最优的. 在所有采用 N 次函数值计算和划界求单峰函数的极小值点的非随机搜索方法中，对于在 N 次

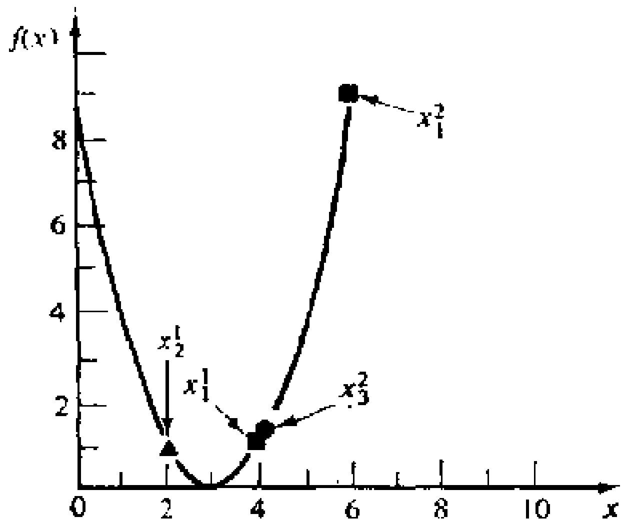


图 8.6 进行 $N=4$ 次函数值计算的 Fibonacci 搜索法

函数值计算以后仍保留的、并包含要寻找的极小值点的最大可能的区间来说，Fibonacci 法将这个区间的长度减少到最小. 换言之，最优搜索方法使初始区间同 N 次函数值计算以后仍保留的并包含极小值点的可能的最终区间的长度之比达到最大. 根据 Avriel 和 Wilde[1]，我们介绍被称为对称 Fibonacci 法的一个最优性证明，这个方法与 Kiefer 方法稍有差异，并且还保证，在最后的迭代中，单峰函数在相互距离为 ε、并且对称地位于区间 $[l_{N-1}, r_{N-1}]$ 中点两侧的两点上计算函数值. 这一修正比 Kiefer 方法在 ε 的量级上略有改进，而当 $\varepsilon \to 0$ 时，这两种方法就变成相同的了.

我们把在实直线的闭区间 $[l_1, r_1]$ 上的一个搜索策略 $S(N, \varepsilon)$ 定义为如下计划：在这个区间的 N 个不同点 $x^1, x^2, \cdots, x^N$ 上计算函数值，这里 x^{k+1} 的位置依赖于单峰函数 f 在 $x^1, \cdots, x^k$ 的值，并且对于一切 $j \neq k$，使得

$$|x^j - x^k| \geqslant \varepsilon, \tag{8.66}$$

这里 $\varepsilon \leqslant \frac{1}{2}$. 这个计划在初始区间逐次缩减到单位长度的最终区间 $[l_N, r_N]$ 后结束，最终区间包含 $[l_1, r_1]$ 上 f 的极小值点 x^*. 这里要注意，选取最终区间具有单位长度只是为了简化记号和证明. 同理，我们假设 $l_1=0$. 按适当的标度，初始区间通常总能表成这种形式.

带有 N 个函数值计算的**对称 Fibonacci** 搜索法是一种关于初始区间$[0, r_{1,N}^F]$的搜索策略，这里右端点为

$$r_{1,N}^F = F_{N+1} - F_{N-1}\varepsilon. \tag{8.67}$$

假定第 k 次函数值计算已经完成，令 x^{k*} 是这样的数值使得

$$f(x^{k*}) = \min\{f(x^1), \cdots, f(x^k)\}. \tag{8.68}$$

规定

$$l_k^F = \max[\{x^j : x^j < x^{k*}, \ j=1, \cdots, k\} \cup \{0\}], \tag{8.69}$$

$$r_k^F = \min[\{x^j : x^j > x^{k*}, \ j=1, \cdots, k\} \cup \{r_{1,N}^F\}]. \tag{8.70}$$

因为 f 是单峰的，显然 x^* 必须属于$[l_k^F, r_k^F]$．要计算函数值的点由下面关系式给出：

$$x^1 = F_N - F_{N-2}\varepsilon = r_{1,N-1}^F, \tag{8.71}$$

$$x^{k+1} = l_k^F + r_k^F - x^{k*}, \quad k=1, \cdots, N-1. \tag{8.72}$$

注意，由(8.67)，点 x^1 恰为用 $N-1$ 次函数值计算的 Fibonacci 搜索中的 $r_{1,N-1}^F$.

首先我们将证明对称的 Fibonacci 法的确是如上定义的搜索策略.

引理 8.6

当 $\varepsilon \leqslant \dfrac{1}{2}$ 时，如上所述的对称的 Fibonacci 搜索对于任何单峰函数来说是在$[0, r_1^F]$上的 $S(N, \varepsilon)$策略.

【证明】 我们必须指出，N 次函数值计算以后，得到单位长度的最终区间$[l_N^F, r_N^F]$，同时没有相邻点位置比 ε 更接近．证明采用对 N 的归纳法．对于 $N=1$，引理显然成立，因为由(8.67)至(8.71)，得到 $l_1^F=0$, $r_{1,1}^F=1$.

对于 $N=2$, $r_{1,2}^F=2-\varepsilon$ 且 $x^1=1$，这蕴涵 $x^2=0+2-\varepsilon-1=1-\varepsilon$. 因此$|x^2-x^1|=\varepsilon$，同时，对于任一单峰函数，或者 $x^{2*}=x^1$，或者 $x^{2*}=x^2$，而在这两种情况下，均有 $r_2^F-l_2^F=1$.

现在假定对 $N\geqslant 2$ 引理成立，那么对于 $N+1$ 次函数值计算来说，有

$$r_{1,N+1}^F = F_{N+2} - F_N\varepsilon, \tag{8.73}$$

$$x^1 = F_{N+1} - F_{N-1}\varepsilon = r^F_{1,N}, \tag{8.74}$$

$$x^2 = F_{N+2} - F_N\varepsilon - F_{N+1} + F_{N-1}\varepsilon$$

$$= F_N - F_{N-2}\varepsilon = r^F_{1,N-1}, \tag{8.75}$$

注意，$x^1 - x^2 = r^F_{1,N} - r^F_{1,N-1} \geqslant 1 - \varepsilon \geqslant \varepsilon$. 如果 $x^{2*} = x^2$，那么 $l^F_2 = 0$ 及 $r^F_2 = x^1 = r^F_{1,N}$. 因此 $0 \leqslant x^* \leqslant r^F_{1,N}$，且在 $r^F_{1,N-1}$ 作了一次函数值计算，由归纳法假设，引理成立. 如果 $x^{2*} = x^1$，那么 $l^F_2 = r^F_{1,N-1}$ 和 $r^F_2 = r^F_{1,N+1}$. 通过简单的变量代换

$$\tilde{x} = r^F_{1,N+1} - x, \tag{8.76}$$

得到 $0 \leqslant \tilde{x}^* \leqslant (r^F_{1,N+1} - r^F_{1,N-1}) = r^F_{1,N}$，且在 $r^F_{1,N-1}$ 作了一次函数值计算，引理还是成立. ▮

在下述定理中我们将证明，在所有搜索策略中，Fibonacci 法达到初始区间右端点的最高值，这相当于说，Fibonacci 法产生初始区间对最终区间的最大长度比值，该最终区间包含着任何单峰函数的极小值点. 然而，一种例外情况必须指出. 前面我们已提到，如果 $f(x^{k+1}) = f(x^{k*})$，那就能够推出 x^* 必定落在 x^{k+1} 与 x^{k*} 之间. 这种情形可以要求搜索方法作适当修改，这里我们就不讨论了.

定理 8.7

在初始区间 $[0, r_{1,N}]$ 上所有搜索策略 $S(N, \varepsilon)$ 中，Fibonacci 法 $S_F(N, \varepsilon)$ 产生右端点的最高值，只要对于 $k = 1, \cdots, N-1$ 有 $f(x^{k+1}) \neq f(x^{k*})$.

【证明】 我们还用对 N 的归纳法. 对于 $N=1$，证明是显然的，因为所有 $r_{1,1} = r^F_{1,1} = 1$. 对于 $N=2$，不失一般性，设在任何策略 $S(2, \varepsilon)$ 中 $x^1 > x^2$. 那么 $x^1 \leqslant 1$，$r_{1,2} - x^1 \leqslant 1 - \varepsilon$，因此 $r_{1,2} \leqslant 2 - \varepsilon = r^F_{1,2}$. 现假定，对 $N \geqslant 2$ 定理成立. 那么对于策略 $S(N+1, \varepsilon)$，设 $x^1 > x^2$ 与 $x^{2*} = x^2$. 因此 $l_{2,N+1} = 0$，$r_{2,N+1} = x^1$，我们得到在 $[0, x^1]$ 上的 $S(N, \varepsilon)$ 策略. 由于归纳法假设，有

$$x^1 \leqslant r^F_{1,N}, \tag{8.77}$$

然而，也由于归纳法假设，应有

$$x^2 \leqslant r^F_{1,N-1}, \tag{8.78}$$

这是因为可能出现 $x^* < x^2$，于是 x^* 必须通过 $[0, x^2]$ 上的 $S(N-1, \varepsilon)$ 策略来确定位置. 如果 $x^{2*} = x^1$，那么 $l_{2,N+1} = x^2$ 和 $r_{2,N+1} = r_{1,N+1}$. 由归纳法假设，有

$$r_{1,N+1} - x^2 \leqslant r_{1,N}^F. \tag{8.79}$$

将(8.78)与(8.79)相加，得到原来所断言的结果如下：

$$r_{1,N+1} \leqslant r_{1,N}^F + r_{1,N-1}^F = r_{1,N+1}^F. \tag{8.80}$$

关系式(8.77)至(8.79)对于 Fibonacci 搜索来说恰成为等式. 另外，用相似于对(8.78)的讨论，我们可得出

$$r_{1,N+1} - x^1 \leqslant r_{1,N-1}^F, \tag{8.81}$$

根据(8.77)和(8.81)可知，这是唯一的一个办法使(8.80)成为等式. 因此，Fibonacci 法是得到 $r_{1,N}$ 最大值的唯一办法. ▌

Fibonacci 法的缺点之一是，在开始搜索之前计算函数值的次数 N 必须预先知道. 这个要求对于所谓**黄金分割法**就不是必需的，它是 Fibonacci 搜索的一个很好近似. 可以证明，有

$$\lim_{N\to\infty} \frac{F_{N-1}}{F_N} = \frac{1}{\tau} = \frac{\sqrt{5}-1}{2} \cong 0.618, \tag{8.82}$$

因此对(8.52)给出的 x_k^2 的位置，可以用下式来近似，

$$x_k^2 \cong l_k + \frac{1}{\tau}(r_k - l_k). \tag{8.83}$$

这样，黄金分割法把计算函数值的点放在下述位置：

$$x_{kG}^1 = l_k + \frac{\tau-1}{\tau}(r_k - l_k), \tag{8.84}$$

$$x_{kG}^2 = l_k + \frac{1}{\tau}(r_k - l_k). \tag{8.85}$$

这里下标 G 表示黄金分割法所取的点，而其他记号保留不变.

从(8.84)和(8.85)看出，黄金分割法同样将有关点对称地放置. 假定 $\varepsilon \to 0$，从(8.56)可以看出，由 N 次函数值计算的 Fibonacci 法，用因子 $1/F_{N+1}$ 缩减包含 f 的极小值点的区间. 这里要求读者证明黄金分割法用因子 $1/(\tau)^{N-1}$ 缩减这个区间. 还可以证明

$$\lim_{N\to\infty}\frac{F_{N+1}}{(\tau)^{N-1}}=\frac{(\tau)^2}{\sqrt{5}}\cong 1.17. \tag{8.86}$$

因此对于大的 N（和小的 ε），黄金分割法所产生的最终区间比 Fibonacci 法要长 17%．或许，两个方法的更为清楚的比较可用下述事实来阐明：用 $N+1$ 次函数值计算的黄金分割法通常能产生的最终区间，其长度与用 N 次函数值计算的 Fibonacci 法得到的区间长度相同或较小．

"黄金分割"的名字，来源于用特种方法分割线段或矩形的古典问题，有关黄金分割法的更完整的讨论以及它与 Fibonacci 法的关系，读者可参考 Wilde[15] 和 Wilde、Beightler[16]的著作．

8.4 最优分批搜索法和黄金分批搜索法

在 8.3 节中所述的方法是基于有顺序的搜索．换言之，需要极小化的函数在某一点上计算函数值，而下一步计算函数值的点的选择依赖于前面的结果．在[15]中还得出另一个效率低得多的方法（相对于包含极小值点位置的初始区间对最终区间的缩减比值而言），在这个方法中，函数值计算常是同时完成的．在某些实际情况下，这样的搜索过程可能是有用的．例如，要寻找极小的函数值计算是一项复杂而费时的工作，但同时测定这些函数值却是可能的．

同时性方法已经与 Fibonacci 法和黄金分割法相结合，成为**分批搜索**技术，在其中要完成一系列同时计算函数值的"批"．要获得比目前运行中的计算机更大型更快速的计算机，一种可能的途径是给它们装备平行处理单元[13]使得算术运算可以同时完成．分批搜索方法是为这种类型计算机导出的第一种最优化技术．其思想是，在平行处理单元中同时完成一批函数值计算，缩减包含极小值点的区间，然后又在留下的区间中完成新的一批函数值计算，以此类推．这些方法由 Avriel 和 Wilde[2, 3] 发展，是基于广义 Fibonacci 数和黄金分割．这些方法的最优性，其意义相当于 Fibonacci 方法和黄金分割法的情况，能类似于引理 8.6 和定理

8.7 那样证明．遵循不同线索的证明，可见 Karp、Miranker 的著作[8]．这里我们只概要叙述算法，其最优性的证明能在前面的参考文章中找到．

分批搜索方法由 n 个有序的批所组成．每个批包含 ξ 个函数值的同时计算．初始区间为 $[0, r^B_{\xi,n}]$（其中上标 B 表示"分批搜索"）．而在 $N=\xi n$ 次函数值计算以后，它被缩减到单位长度．再次令 $\varepsilon\leqslant\frac{1}{2}$．

根据每批中函数值同时计算的数目是 $\xi=2k-1$（奇数）或是 $\xi=2k$（偶数），我们有两个不同的策略，概述如下：

(a) 用于 $\xi=2k-1$（奇数）的分批搜索法

利用下述递推关系定义广义 Fibonacci 序列：

$$A(k, h+2)=k[A(k, h+1)+A(k, h)], \tag{8.87}$$

其中 $A(k, 0)=0$, $A(k, 1)=1$．注意该序列对于负整数 h 也有定义．对于 $k=1$，关系式 (8.87) 就成为熟知的 Fibonacci 序列．在奇数分批搜索中初始区间的右端点由下式给出：

$$r^B_{2k-1,n}=A(k, n+1)-kA(k, n-1)\varepsilon. \tag{8.88}$$

由于要求极小化的函数假定是单峰的，所以在第 i 批函数值计算后，我们就得到一个缩减的包含点 x^{i*} 的区间，在 x^{i*} 该函数具有该批所有点上的最低值（计算了 $2k-1$ 个点，再加前一批留下一个点）．尽可能利用原来的表示法．对 $i=1, 2, \cdots, n$，令 $x^0_i=l^B_{i-1}$，并令 $x^1_i, x^2_i, \cdots, x^{2k}_i$ 是第 i 批中函数值同时计算的 $(2k-1)$ 个点的位置及前已完成计算的 $x^{(i-1)*}$ 的位置，则

$$x^1_i=l^B_{i-1}+A(k, n-i+1)-kA(k, n-i-1)\varepsilon, \tag{8.89}$$

$$x^{j+2}_i=x^j_i+\frac{1}{k}A(k, n-i+2)-A(k, n-i)\varepsilon,$$

$$j=0, \cdots, 2k-2. \tag{8.90}$$

(b) 用于 $\xi=2k$（偶数）的分批搜索法

初始区间的右端点由下式给出：

$$r^B_{2k,n}=(k+1)^n-\varepsilon. \tag{8.91}$$

同样令 $x_i^0=l_{i-1}^B$. 那么第 i 批的 $2k$ 个函数值计算及前已完成计算的 $x^{(i-1)*}$ 放置如下:

$$x_i^{j+2}=x_i^j+(k+1)^{n-i},\quad j=0,\ \cdots,\ 2k-1, \tag{8.92}$$

$$x_1^1=(k+1)^{n-1}-\varepsilon. \tag{8.93}$$

对于 $i=2,\ \cdots,\ n$, 或者 $x_i^1=x^{(i-1)*}=l_{i-1}^B+\varepsilon$, 或者 $x_i^{2k+1}=x^{(i-1)*}=r_{i-1}^B-\varepsilon$. 故而 $x_i^2,\ x_i^4,\ \cdots,\ x_i^{2k}$ 可以利用(8.92)从 x_i^0 开始递推地找到, 而利用(8.92)以 $x^{(i-1)*}$ 作为起始点, 定出 $x_i^1,\ x_i^3,\ \cdots,\ x^{2k+1}$ 的位置.

例 8.4.1

假设我们要划定区间[0, 40]上定义的单峰函数的极小值点, 使最终区间为单位长度, 而 $\varepsilon=0.5$. 由于实际原因, 只能分三批计算函数值, 虽然每一批中函数值同时计算的数目不加限制.

首先必须确定 ξ, 即每一批中函数值同时计算的最小数目, 使得在做完 3 批之后得到一个单位区间. 利用(8.88)和(8.91), 其中 $n=3$, 我们就得到 $r_{1,3}^B=2.5$, $r_{2,3}^B=7.5$, $r_{3,3}^B=14$, $r_{4,3}^B=26.5$ 以及 $r_{5,3}^B=40.5$. 在完成每一批的 $\xi=5(k=3)$ 个函数值计算之后, 我们就确保初始区间缩减 40 倍. 第一批中计算函数值的点由(8.89)和(8.90)确定如下:

$$\left.\begin{aligned}
x_1^1&=12-(3)(1)(0.5)=10.5,\\
x_1^2&=\frac{1}{3}(45)-(3)(0.5)=13.5,\\
x_1^3&=10.5+\frac{1}{3}(45)-(3)(0.5)=24.0,\\
x_1^4&=13.5+\frac{1}{3}(45)-(3)(0.5)=27.0,\\
x_1^5&=24+\frac{1}{3}(45)-(3)(0.5)=37.5.
\end{aligned}\right\} \tag{8.94}$$

图 8.7 表示出第一批中点上的函数值. 可以看到, 在第一批之后保留下来的包含极小值点的区间是[13.5, 27.0].

现在, 在第二批的下列点上(包括了保留下来的点 24.0)计算函数值.

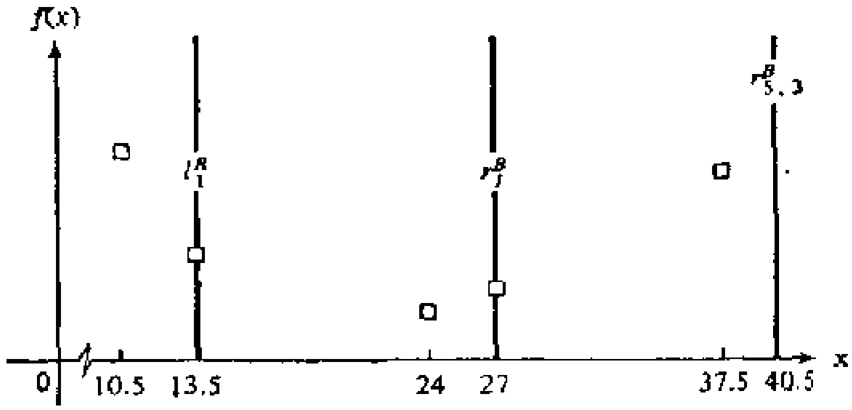


图 8.7　函数值计算的第一批

$$\left.\begin{aligned}
x_2^1&=13.5+3-(3)(0)(0.5)=16.5,\\
x_2^2&=13.5+\frac{1}{3}(12)-(1)(0.5)=17.0,\\
x_2^3&=16.5+\frac{1}{3}(12)-(1)(0.5)=20.0,\\
x_2^4&=17.0+\frac{1}{3}(12)-(1)(0.5)=20.5,\\
x_2^5&=20.0+\frac{1}{3}(12)-(1)(0.5)=23.5,\\
x_2^6&=20.5+\frac{1}{3}(12)-(1)(0.5)=24.0.
\end{aligned}\right\}\tag{8.95}$$

在第二批函数值计算后的状况如图 8.8 所示，保留下来的区间是 [23.5, 27.0]，保留下来的点仍是 24.0.

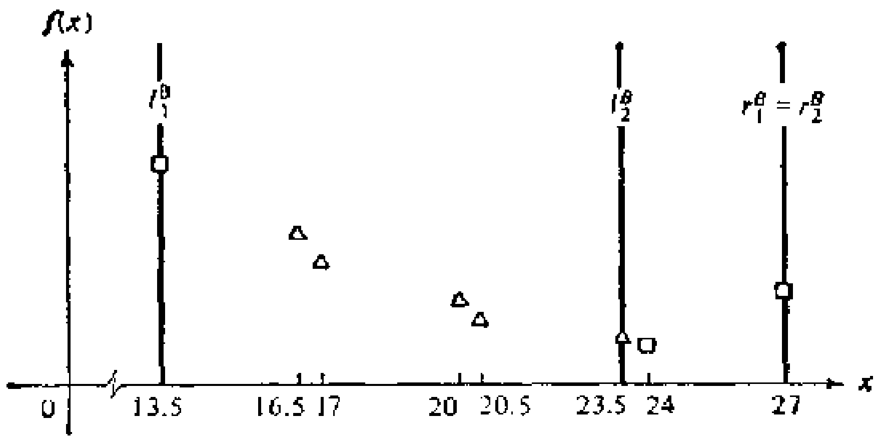


图 8.8　函数值计算的第二批

读者可以利用(8.89)和(8.90)验证，第三批中计算函数值的点以 $\frac{1}{2}$ 单位长的距离间隔地放置，如图8.9所示. 最终区间[24.5, 25.5]为单位长度. ▌

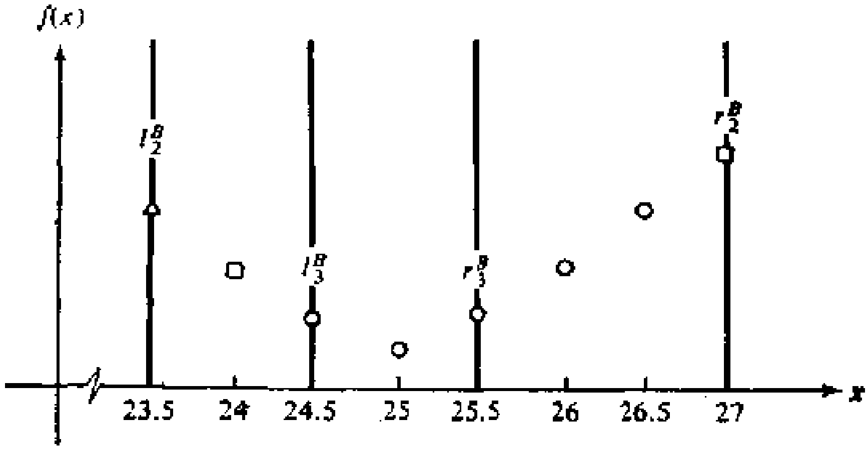


图 8.9 函数值计算的第三批

在将要结束关于最优分批搜索法的时候，我们给出一些例子来说明它的效率，该效率以缩减到单位长度最终区间的初始区间的长度为依据. 在表 8.1 中列出了 $\varepsilon=\frac{1}{2}$ 及分批数 2、5、8 情况下的初始区间长度，也就是 $r_{\xi,n}^B$ 的值，将它作为每一批中同时计算函数值次数的一个函数. 这个表说明，对于给定的 ξ，初始区间长度随着分批数目而巨大地增加，又对于给定分批数，它也随着每批中函数值同时计算的数目的增加而可观地增加.

表 8.1 初始区间长度

ξ	$r_{\xi,2}^B$	$r_{\xi,5}^B$	$r_{\xi,8}^B$
1	1.5	6.5	27.5
2	3.5	31.5	255.5
3	5.0	104.0	2110.0
4	8.5	242.0	6560.5
5	10.5	580.5	31630.5
6	15.5	1023.5	65535.5

在给定函数值计算总数时，观察不同的分配方案下分批搜索法的效率也是颇为有趣的. 表 8.2 比较了在各种不同的分批方案中安排 12 次函数值计算所相应的初始区间长度. 这里，仍取 $\varepsilon=\frac{1}{2}$.

表 8.2　12 次函数值计算的初始区间长度

ξ	n	$\tau_{\xi,n}^{B}$
1	12	188.5
2	6	63.5
3	4	38.0
4	3	26.5
6	2	15.5
12	1	6.5

可以看出，对给定的函数值计算总数，Fibonacci 搜索法具有可以缩减到单位长度的最大初始区间，因为它利用了函数值计算所提供的所有有用信息，而没有任何一次函数值计算被分批搜索法的同时性特征所浪费掉.

最优分批搜索法的进一步推广是由 Beamer 和 Wilde[5]作出，他们容许函数值同时计算的数目可以逐批变更. 进一步细节读者可参考[5].

最优分批搜索法与 Fibonacci 搜索法有类似缺点；即搜索中所用到的分批数必须是预先已知的. 正如黄金分割法是 Fibonacci 法的良好近似，下面要描述的黄金分批搜索法甚至是最优分批搜索法更为良好的近似.

Avriel、Wilde[3] 已经证明，

$$\lim_{n\to+\infty}\frac{A(k,\ n+1)}{A(k,\ n)}=\frac{k+[(k)^2+4k]^{1/2}}{2}=\tau_k, \tag{8.96}$$

即对 $k=1,\ 2,\ \cdots$，数 τ_k 是下述方程的解：

$$(t)^2-kt-k=0. \tag{8.97}$$

对于 $k=1$，τ_k 的值等于在黄金分割法中用到的 τ. 对于 $k=2$，τ_k 的值近似为 2.732，$\tau_3=3.791$ 等等.

首先考虑每批中函数值计算奇数次，当 $\varepsilon\to 0$ 时，对最优分批搜索法，我们得到

$$r^{B}_{2k-1,n}\cong A(k,\ n+1), \tag{8.98}$$

$$x^1_i\cong l^B_{i-1}+A(k,\ n-i+1) \tag{8.99}$$

以及

$$x^{j+2}_i\cong x^j_i+\frac{1}{k}A(k,\ n-i+2),\quad j=0,\ \cdots,\ 2k-2. \tag{8.100}$$

对于大的 n，我们从(8.96)得到

$$\frac{A(k,\ n)}{A(k,\ n-i)}\cong(\tau_k)^i. \tag{8.101}$$

现在，假定黄金分批搜索的初始区间等于某个最优分批搜索的初始区间，这最优分批搜索具有 n 批，每批各计算 $2k-1$ 次函数值且 $\varepsilon=0$. 即

$$r^{GB}_{2k-1}=r^{B}_{2k-1,n}, \tag{8.102}$$

这里的上标 GB 表示"黄金分批". 然后利用(8.98)至(8.102)，按照黄金分批搜索法，我们在下列点计算函数值：

$$x^1_i=l^{GB}_{i-1}+\left(\frac{1}{\tau_k}\right)^i r^{GB}_{2k-1}, \tag{8.103}$$

$$x^{j+2}_i=x^j_i+\frac{1}{k}(\tau_k)^{1-i}r^{GB}_{2k-1},\quad j=0,\ \cdots,\ 2k-2. \tag{8.104}$$

各计算 $2k-1$ 次函数值的 n 批以后，包含单峰函数极小值点的最终区间之长度由 $(1/k)(\tau_k)^{1-n}r^{GB}_{2k-1}$ 给出.

再考虑，每批中函数值计算偶数次，这时，黄金分批搜索可通过重新安排最优偶数分批搜索的相应关系式而得到. 于是

$$x^1_1=\frac{r^{GB}_{2k}-k\varepsilon}{k+1}, \tag{8.105}$$

$$x^{j+2}_i=x^j_i+\frac{r^{GB}_{2k}+\varepsilon}{k+1},\quad j=0,\ \cdots,\ 2k-1. \tag{8.106}$$

这些表达式仅包含初始区间 $[0,\ r^{GB}_{2k}]$ 和间隔 ε，而与分批数无关. 在每批 $2k$ 个函数值同时计算的 n 批之后，最终区间长度由 $(r^{GB}_{2k}+\varepsilon)/(k+1)^n$ 给出.

例 8.4.2

假设我们要寻求区间[0, 45]上单峰函数的极小值点. 容易算出对于 $\varepsilon\to 0$, 有

$$r_{5,3}^{B}\cong A(3,\ 4)=45. \tag{8.107}$$

从而，利用最优分批搜索法确定的每批有5个函数值同时计算的3批之后，可以把初始区间缩减到单位长的最终区间. 让我们在同一区间上进行5个函数值同时计算的黄金分批搜索. 表8.3给出了在第一批中函数值计算的位置，它由(8.103)和(8.104)给出. 括号内的数字指出相应的最优分批搜索所确定的函数值计算的位置.

表 8.3 第一批中函数值计算的位置

j	1	2	3	4	5
x_j^1	11.87(12)	15(15)	26.87(27)	30(30)	41.87(42)

用黄金分批搜索所确定的第三批函数值计算之后，余下区间长度为 $\frac{1}{3}(3.791)^{-2}(45)=1.0436$. 即与最优分批搜索法所能得到的相应的单位长度最终区间相比，这区间的长度仅仅超过4.36%. ∎

通常，在每批中函数值同时计算的数目越多，那么用黄金分批搜索来近似最优分批方法也就越好.

练 习

8.A. 写出实现 Newton 法和割线法的简短计算程序. 用它们找出简单方程的解. 计算每次迭代的误差，并且将得到的试验收敛速率与理论收敛速率进行比较.

8.B. 证明：不等式(8.34)蕴涵(8.32)中二次项系数 c 为正数，且关于 ϕ 预测的逗留点确实是极小值点.

8.C. 假定要求极小的函数的一阶导数处处是已知的，试修正二次方法和划界方法.

8.D. 设 f 是 R 上实函数，又设 $x^0\in R^n$，$z\in R^n$，$\theta\in R$. 定义

$$F(\theta)=f(x^0+\theta z), \tag{8.108}$$

并假定我们寻求 F 的极小，即经过点 x^0 在方向 z 上函数 f 的极小. 令 $x^0+\theta_1 z$，$x^0+\theta_2 z$ 和 $x^0+\theta_3 z$ 三点上 f 已算得数值. 证明：利用二次逼近方法求得的极小值点为 $x^0+\theta^* z$，此处[11]

$$\theta^*=\frac{[(\theta_2)^2-(\theta_3)^2]F(\theta_1)+[(\theta_3)^2-(\theta_1)^2]F(\theta_2)+[(\theta_1)^2-(\theta_2)^2]F(\theta_3)}{2[(\theta_2-\theta_3)F(\theta_1)+(\theta_3-\theta_1)F(\theta_2)+(\theta_1-\theta_2)F(\theta_3)]}, \tag{8.109}$$

并且，如果

$$\frac{(\theta_2-\theta_3)F(\theta_1)+(\theta_3-\theta_1)F(\theta_2)+(\theta_1-\theta_2)F(\theta_3)}{(\theta_2-\theta_3)(\theta_3-\theta_1)(\theta_1-\theta_2)}<0, \tag{8.110}$$

那么它确实是通过上面三点的抛物线的极小值点.

8.E. 划界方法和二次逼近方法的下述形式见 Davies、Swann、Campey[14]. 在某一初始点 $x^0\in R$，计算函数 f 的数值，取预定步长 $h>0$，然后在 x^0+h 计算函数 f 的数值. 如果

$$f(x^0)\geqslant f(x^0+h), \tag{8.111}$$

将步长加倍，并且在函数值减少的方向计算 f 的后续值. 重复这过程直至得到函数值的增加.

如果

$$f(x^0)<f(x^0+h), \tag{8.112}$$

于是计算 $f(x^0-2h)$，$f(x^0-4h)$ 等等直至得到函数值的增加，最后三点形成包含 f 的极小值点的界限. 于是在用此法得到的最后两点之间的中点上计算函数值. 这个过程产生按某个等距离 s 放置的 4 个点，抛弃离函数值最低点为最远的端点，而保留三个等间隔点以便进行二次逼近. 设 $x^1<x^2<x^3$ 是这三个点，并设 f^i 是相应的函数值. 证明：用二次逼近方法所预测的极小值点 x^4 由下式给出

$$x^4=x^2+\frac{s(f^1-f^3)}{2(f^1-2f^2+f^3)}, \tag{8.113}$$

然后，用较小的步长重复进行划界过程，x^4 能够相应于通过三点的抛物线的极大值点吗？画出这个方法的框图.

8.F. 详细描述求 R 上处处有定义的函数极小的一种搜索法. 这种方法以带有外推法的二次逼近为基础(即不用划界)仅利用函数值计算，并画出该方法的框图.

8.G. 验证三次逼近方法的关系式(8.44)至(8.47).

8.H. 证明通过 N 次函数值计算的黄金分割法，最终区间对初始区间的长度比率由 $1/(\tau)^N$ 给出. 用下述关系

$$F_k = \frac{(\tau)^k - (-\tau)^{-k}}{\sqrt{5}}, \tag{8.114}$$

证明(8.86)式.

8.I. 利用最优分批搜索法的公式导出对称的 Fibonacci 搜索法 ($\xi=1$), 并且同本章给出的公式比较.

8.J. 利用最优分批搜索法的公式,对任意的奇数次或偶数次函数计算，导出同时性搜索过程($n=1$). 并对奇数次情形和偶数次情形比较该方法的效率.

参 考 文 献

1. AVRIEL, M., and D. J. WILDE, "Optimality Proof for the Symmetric Fibonacci Search Technique," *Fibonacci Quarterly*, **4**, 265–269 (1966).
2. AVRIEL, M., and D. J. WILDE, "Optimal Search for a Maximum with Sequences of Simultaneous Function Evaluations," *Management Science*, **12**, 722–731 (1966).
3. AVRIEL, M., and D. J. WILDE, "Golden Block Search for the Maximum of Unimodal Functions," *Management Science*, **14**, 307–319 (1968).
4. BARTLE, R. G., *The Elements of Real Analysis*, 2nd ed., John Wiley & Sons, New York, 1976.
5. BEAMER, J. H., and D. J. WILDE, "Minimax Optimization of Unimodal Functions by Variable Block Search," *Management Science*, **16**, 529–541 (1970).
6. DAVIDON, W. C., "Variable Metric Method for Minimization," AEC Research and Development Report ANL-5990 (Rev.), November 1959.
7. GOLDSTEIN, A. A., *Constructive Real Analysis*, Harper & Row, New York, 1967.
8. KARP, R. M., and W. L. MIRANKER, "Parallel Minimax Search for a Maximum," *J. Combinatorial Theory*, **4**, 19–35 (1968).
9. KIEFER, J., "Sequential Minimax Search for a Maximum," *Proc. Am. Math. Soc.*, **4**, 502–506 (1953).
10. OSTROWSKI, A. M., *Solution of Equations and Systems of Equations*, 2nd ed., Academic Press, New York, 1966.
11. POWELL, M. J. D., "An Efficient Method for Finding the Minimum of a Function of Several Variables without Calculating Derivatives," *Computer J.*, **7**, 155–162 (1964).
12. ROSENBROCK, H. H., "An Automatic Method for Finding the Greatest or Least Value of a Function," *Computer J.*, **3**, 175–184 (1960).

13. SLOTNICK, D. L., "The Fastest Computer," *Scientific American*, **224**, 2, 76–87 (1971).

14. SWANN, W. H., "Report on the Development of a New Direct Search Method of Optimization," I.C.I. Central Instrument Lab. Research Note 64/3, London, 1964.

15. WILDE, D. J., *Optimum Seeking Methods*, Prentice-Hall, Englewood Cliffs, N.J., 1964.

16. WILDE, D. J., and C. S. BEIGHTLER, *Foundations of Optimization*, Prentice-Hall, Englewood Cliffs, N.J., 1967.

第 9 章

不用导数的多维无约束最优化：经验方法与共轭方向法

对一个函数了解得越多，我们就可以导出寻找其极值点的更好的或更有效的算法；但是另一方面，为了得到该函数的更多的信息，却又可能需要大量的计算工作．例如，如果一个函数是凸的，则它的每一个逗留点必然对应到一个整体极小，但是为确定凸性所需的计算量却可能是很大的．类似地，如果 R^n 上的一个函数的一阶导数与二阶导数可供使用，这将大大有利于确定其极值点的位置，但是为计算其导数（特别是对于大的 n）所需的算术运算次数有时是这样的大，以致我们宁愿试着去另找一个不用导数的算法．在许多实际问题中，这些导数简直是不能使用的，有时甚至连要极小化的函数的解析表达式也不能找到．这些考虑说明，有必要提出一些只需计算函数值的无约束多维最优化算法．

在以后几章中将讨论一些更有效的算法，其中不仅要求对函数进行计算，而且要求对其导数进行计算．这些算法中，有的已被修改，使导数无须直接计算而只是用其函数值来逼近．等讲到有关算法时，我们将提到这些修改方案．在本章前三节的每一节中我们将讲述一个经验方法，这些方法代表了无约束最优化的早期成就．虽然今天有着更先进的方法可供使用，但经验方法及其基本原则的简明性使得它们在某些情况下仍然是有用的．本章的其余部分处理利用共轭方向的方法．在此处引入共轭方向的概念是为了理解不用导数的 Powell 方法，该方法一般认为优于经验方法．

9.1 单纯形法

长期以来人们就认识到，为了寻找几个变量的函数的极小值

点，只用最简单的概念——譬如在 R^n 上产生一个网格，并在此网格的每一点上计算函数值；或者用随机移动来搜索极小值点——是很低效的．改进的经验方法在60年代初期就出现了，其第一个就是所谓单纯形法，现在我们就来介绍它．这里顺便提醒一下，无约束极小化的单纯形法决不应同线性规划中的单纯形法混淆起来，虽然二者命名的出发点是相同的．一个单纯形是 R^n 中 $n+1$ 个点的一个凸包，例如在 R 中的一个线段，在 R^2 中的一个三角形等等．无约束极小化的单纯形法是 Spendley、Hext、Himsworth[15] 设计的，随后为 Nelder、Mead[7] 所改进．

考虑实函数 $f(x)(x\in R^n)$ 的极小化，设 $x^0, x^1, \cdots, x^n$ 是 R^n 中的点，由它们构成一个当前的单纯形．x^h 及 x^l 由下式定义：

$$f(x^h)=\max\{f(x^0), f(x^1), \cdots, f(x^n)\}, \tag{9.1}$$

$$f(x^l)=\min\{f(x^0), f(x^1), \cdots, f(x^n)\}. \tag{9.2}$$

用 $\bar{x}$ 记这个单纯形除 x^h 外的所有顶点的形心：

$$\bar{x}=\frac{1}{n}\sum_{i=0}^{n}x^i, \quad x^i\neq x^h. \tag{9.3}$$

这个算法的主要想法是用一个新的、比较好的点来替换当前单纯形中有着最高函数值的顶点 x^h，这点的替换包含三种类型的步骤：反映、扩展和压缩．

在反映步骤中，用下式计算 x^r：

$$x^r=\bar{x}+\alpha(\bar{x}-x^h), \tag{9.4}$$

其中 α 是一个正的常数，称为反映系数．

考虑三种可能的场合：

1．如果 $f(x^l)>f(x^r)$，这就是说，反映步骤已产生一个新的最小值，那么我们就进行一个扩展步骤，计算

$$x^e=\bar{x}+\gamma(x^r-\bar{x}), \tag{9.5}$$

其中 $\gamma>1$ 是一个给定的常数，称为扩展系数．若 $f(x^r)>f(x^e)$，则用 x^e 替换 x^h，并得到一个新的单纯形；但是，若 $f(x^e)\geqslant f(x^r)$，则扩展步骤失败，这时用 x^r 替换 x^h 以构成新的单纯形．

2．如果

$$\max\{f(x^i),\ x^i\neq x^h\}\geqslant f(x^r)\geqslant f(x^l), \tag{9.6}$$

那么就用 x^r 替换 x^h，从而产生一个新的单纯形.

3. 如果

$$f(x^r)>\max_i\{f(x^i),\ x^i\neq x^h\}, \tag{9.7}$$

这时若以 x^r 替换 x^h，将使 x^r 成为新的 x^h. 在这种场合，我们用下式定义一个点 $x^{h'}$,

$$f(x^{h'})=\min\{f(x^h),\ f(x^r)\}, \tag{9.8}$$

并进行一个压缩步骤

$$x^c=\bar{x}+\beta(x^{h'}-\bar{x}), \tag{9.9}$$

其中 $0<\beta<1$，是压缩系数. 一般就以 x^c 替换 x^h 而构成新的单纯形，除非是 $f(x^c)>f(x^{h'})$，而这时则通过下式定义新的点 $\hat{x}^i$ 来替换所有的 x^i,

$$\hat{x}^i=x^i+\frac{1}{2}(x^l-x^i),\quad i=0,\ \cdots,\ n. \tag{9.10}$$

也即把老的单纯形的顶点与具有最低函数值的点的距离取半.

用单纯形法对某些试验函数进行的性能估价指出，选取 $\alpha=1$, $\beta=\frac{1}{2}$, $\gamma=2$ 给出好的结果[7]. 用单纯形法对极小值点的一个典型搜索见图 9.1，其中编号指出相继的单纯形的顶点，虚线表示所作的移动，其中的某些移动因失败而终止. Nelder 和 Mead 提议，算法终结条件为

$$\left\{\frac{1}{n+1}\sum_{i=0}^{n}[f(x^i)-f(\bar{x})]^2\right\}^{1/2}<\varepsilon, \tag{9.11}$$

这里 ε 是某个预先给定的正数.

Spendley、Hext 和 Himsworth 的最初的单纯形法（它们用正规单纯形，没有扩展与压缩步骤）及上述 Nelder 和 Mead 的变型已经在许多问题上进行了成功的试验，但也发现，第一个单纯形所选的尺度与方向对结果有重大影响[14]. 此外有人报道，Nelder 和 Mead 变型（普遍认为它优于最初的方法）对于变量很多的情形，譬如 $n\geqslant10$ 的问题是十分无效的[10].

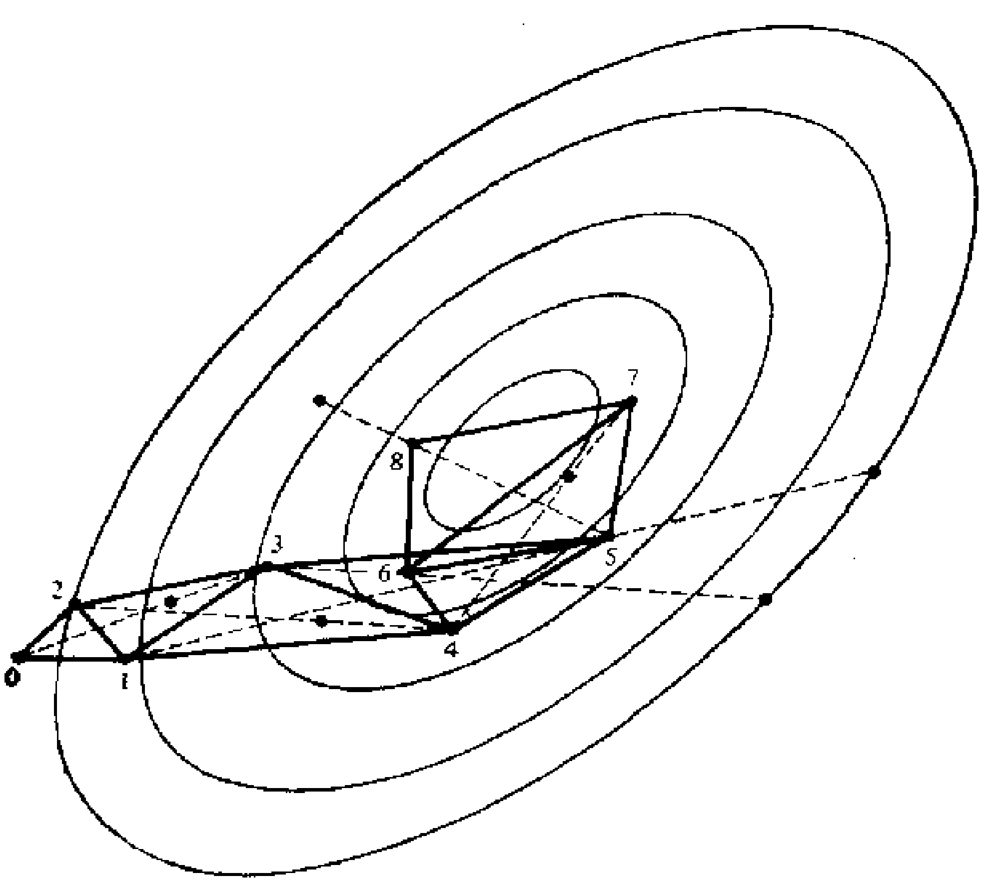


图 9.1 用二次函数说明单纯形法

9.2 模式搜索

这个简单而容易实现的算法属于 Hooke、Jeeves[5]. 仍假定我们是要寻找一个实值函数 $f(x)$ $(x\in R^n)$ 的极小值点. 给定一个初始点 x_B^0(它也记作 t_0^1, 这样做的理由读完下文自明), 我们在 x_B^0 周围进行一系列勘探移动如下: 计算 $f(x_B^0)$ 与 $f(x_B^0+\Delta_1)$, 这里 Δ_j 是一个 n 维向量, 其第 j 个分量是 $d_j>0$, 其余分量为 0; 即

$$\Delta_1=(d_1,\ 0,\ 0,\ \cdots,\ 0)^T, \tag{9.12}$$

其中 d_1 是某个指定的正数(步长). 因此

$$(x_B^0+\Delta_1)=(x_{B_1}^0+d_1,\ x_{B_2}^0,\ \cdots,\ x_{B_n}^0). \tag{9.13}$$

如果 $f(x_B^0)>f(x_B^0+\Delta_1)$, 这个勘探移动是成功的, 我们就取 $t_1^1=x_B^0+\Delta_1$, 并接着去做沿 x_2 轴方向的移动; 否则, 我们反转搜索方向, 计算 $f(x_B^0-\Delta_1)$. 如果 $f(x_B^0-\Delta_1)<f(x_B^0)$, 那么这个勘探移动认为是成功的, 并取 $t_1^1=x_B^0-\Delta_1$; 如果后一移动也失败, 即

$$f(x_B^0-\Delta_1)\geqslant f(x_B^0), \tag{9.14}$$

就取 $t_1^1=t_B^0$.

完成了沿 x_1 轴方向的勘探移动之后，我们接着计算 $f(t_1^1+\Delta_2)$，这里 $\Delta_2=(0,\ d_2,\ 0,\ \cdots,\ 0)^T$. 如果 $f(t_1^1+\Delta_2)<f(t_1^1)$，就取 $t_2^1=t_1^1+\Delta_2$；否则就计算 $f(t_1^1-\Delta_2)$，并把它与 $f(t_1^1)$ 比较. 当 $f(t_1^1-\Delta_2)<f(t_1^1)$时，取 $t_2^1=t_1^1-\Delta_2$，否则取 $t_2^1=t_1^1$. 按此方式，我们沿所有 n 个坐标轴方向作勘探移动并到达一点 t_n^1，称这点为新的基点 x_B^1，并立即沿着"有前途的"方向 $x_B^1-x_B^0$ 移动到

$$t_0^2=x_B^1+(x_B^1-x_B^0), \tag{9.15}$$

这称为**模式移动**. 如果 $t_0^2\neq x_B^1$，就进行一轮新的勘探移动以建立一个新的基点 x_B^2；否则因在 t_0^1 周围进行的勘探移动已全部失败，我们就减小 d_j 并用较小的步长在 t_0^1 周围再作勘探移动. 如果 $f(x_B^2)<f(x_B^1)$，我们作下一个模式移动到

$$t_0^3=x_B^2+(x_B^2-x_B^1) \tag{9.16}$$

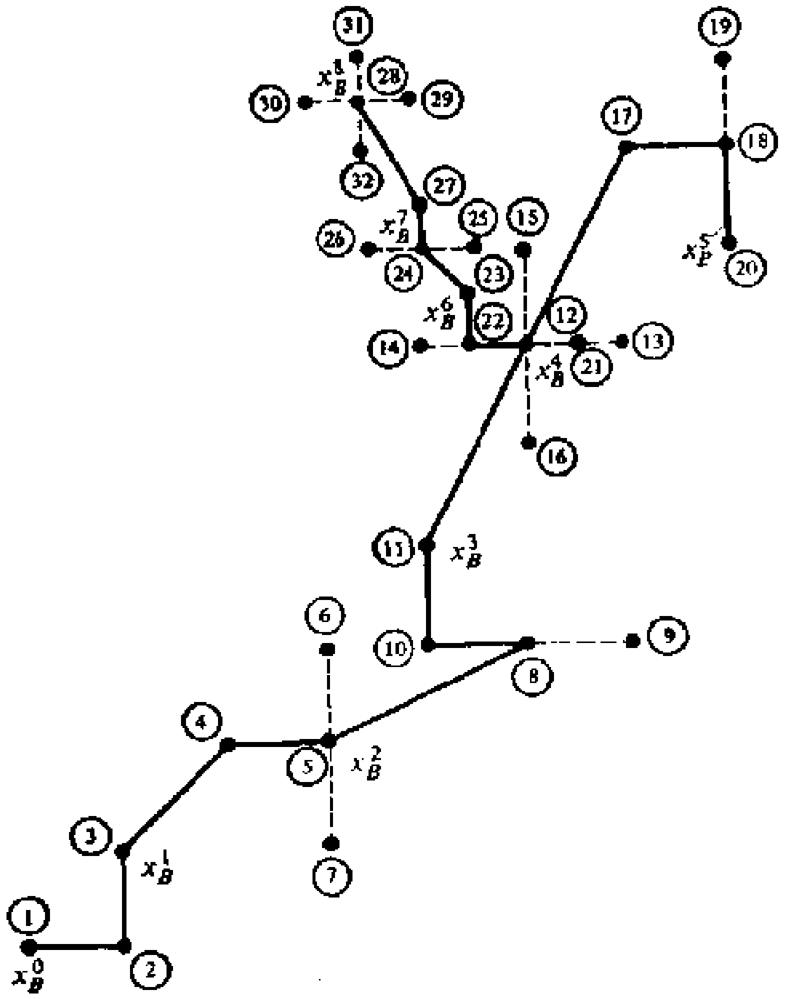
并在 t_0^3 周围重新开始勘探过程.

继续这样做，直至我们达到一个基点 x_B^k，有着$f(x_B^k)\geqslant f(x_B^{k-1})$. 遇到这种场合就回到 x_B^{k-1}，取 $t_0^k=x_B^{k-1}$，并从它重新开始勘探移动. 如果这些勘探移动是成功的，就重新进行另一模式移动与勘探移动；否则，减小 d_j，并用较小的步长在 t_0^k 周围再作勘探移动. 当 d_j 变得小于某些预定值时算法就终结. 图 9.2 说明把这个算法用于某一个双变量函数的情况，圆圈中的数字指明所经过的点的顺序.

为了到达靠近函数极小值点的一个点，虽然模式搜索法可能需要大量的函数值计算，但是它被认为是一个容易编程序而又可靠的方法. 它的主要特征在于顺着"脊"或"谷". 模式移动沿着设想的谷的方向跨出长步，而勘探移动则在一个模式移动已跨越深谷时重新找一条路返回这些深谷.

9.3 旋转方向法

顺着脊或谷的重要性在 Hooke-Jeeves 的模式搜索中得到了明显的反映，同样地，在即将叙述的 Rosenbrock 的旋转方向法[13]中也是如此. 导致提出这两个早期的无约束极小化方法的理由可



所经过的点	函数值
1	48.0
2	46.2
3	42.1
4	44.2
5	40.9
6	41.5
7	41.3
8	32.0
9	35.2
10	30.1
11	27.2
12	12.5
13	15.1
14	12.8
15	13.2
16	17.1
17	16.8
18	16.3
19	17.5
20	15.2
21	13.8
22	12.0
23	10.5
24	9.1
25	12.8
26	10.2
27	8.2
28	5.0
29	6.2
30	7.1
31	8.2
32	7.0

图 9.2 模式搜索用于一个两变量的函数

能是：两个最老的算法——交替方向法与最速下降法，在用于一般函数时性能很差．在交替方向法中，我们从一个任意点出发，沿 x_1 轴方向用前章介绍的某个一维方法搜索一个极小值点．再从此极小值点出发，平行于 x_2 轴进行搜索，按此办理直至所有方向 $x_1, \cdots, x_n$ 被依次搜索过．完成这样一轮之后，整个过程又重新开始．读者不难作出用这种方法只能以极小步长逼近极小值点的那种二维函数的等值线图．最速下降法(将在后面介绍)也有类似的缺陷．这两个方法的主要缺点在于它们没有能力因地制宜地改变自己的搜索方向．

我们来看一看 Rosenbrock 的旋转方向法是如何搜索 R^n 上一

个无约束实值函数 f 的极小值点的. 虽然在 Rosenbrock 原来的著作中曾把这方法推广，使之可用于带有某些约束的极小化，但是在本节中我们只讨论无约束的情形.

从一个给定的初始点 $x_B^0 \in R^n$ 出发，并选取一组 n 个相互正交的方向 $\zeta^1, \cdots, \zeta^n$（通常就选坐标轴的单位向量），接着，沿这些方向进行勘探移动如下：计算 $f(x_B^0)$ 与 $f(x_B^0+\zeta^1)$. 如果 $f(x_B^0+\zeta^1) \leqslant f(x_B^0)$，就称这步为成功，以 $x_B^0+\zeta^1$ 替换 x_B^0，并乘 ζ^1 以数量 $\alpha>1$；如果 $f(x_B^0+\zeta^1)>f(x_B^0)$，则称这步为失败，保持 x_B^0 不变，并乘 ζ^1 以数量 $-\beta$, $0<\beta<1$. 不管哪种场合，我们都进行到下一个方向，并在当前点的周围作同样方式的勘探移动. 对所有 n 个方向都这样做. 与 Hooke-Jeeves 的模式搜索法不同，现在这组沿相互正交方向的勘探移动一直重复进行，直到在所有方向都至少出现一次成功并接着一次失败为止. 算法的这一部分称为勘探阶段. 一个由 8 次函数值计算构成的典型的勘探阶段说明于图 9.3 中，其中圆圈里的数字对应于依次计算函数值的点. 粗线与虚线分别指明成功与失败的移动.

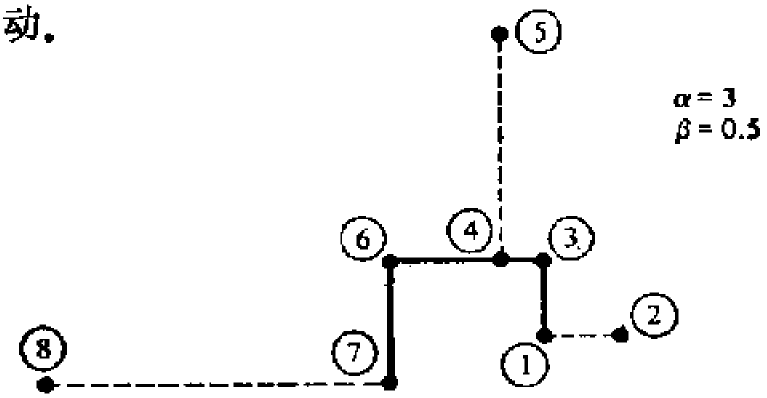


图 9.3　旋转方向法的一个勘探阶段

完成了一个勘探阶段之后，重新构造一组新的正交方向. 以 α_j, $j=1, \cdots, n$, 记勘探阶段中沿着第 j 个当前方向移动的净距离. 令

$$
\begin{aligned}
u^1 &= \alpha_1\zeta^1+\alpha_2\zeta^2+\cdots+\alpha_n\zeta^n,\\
u^2 &= \qquad\quad \alpha_2\zeta^2+\cdots+\alpha_n\zeta^n,\\
&\vdots \qquad\qquad\qquad\qquad \vdots\\
u^n &= \qquad\qquad\qquad\qquad\ \ \alpha_n\zeta^n.
\end{aligned}
\tag{9.17}
$$

注意到 u^1 是连结初始点与勘探阶段终点的向量，它被设想是“有前途”的方向。接着，通过所谓 Gram-Schmidt 正交化过程[8] 找得一组新的正交向量。设(9.17)中所有的 α_j 都不等于 0，那么向量 u^j，$j=1, \cdots, n$，是线性无关的。记

$$w^1=u^1, \tag{9.18}$$

$$\hat{\zeta}^1=\frac{w^1}{\|w^1\|}, \tag{9.19}$$

$$w^2=u^2-[(u^2)^T\hat{\zeta}^1]\hat{\zeta}^1, \tag{9.20}$$

$$\hat{\zeta}^2=\frac{w^2}{\|w^2\|}, \tag{9.21}$$

一般地，

$$w^j=u^j-\sum_{k=1}^{j-1}[(u^j)^T\hat{\zeta}^k]\hat{\zeta}^k, \quad j=1, \cdots, n, \tag{9.22}$$

$$\hat{\zeta}^j=\frac{w^j}{\|w^j\|}, \quad j=1, \cdots, n. \tag{9.23}$$

这样得到的 $\hat{\zeta}$ 为一个新的勘探阶段提供了一组新的相互正交的方向。第一个搜索方向 $\hat{\zeta}^1$ 将通过前一个勘探阶段的起点与终点，新方向的步长一般选得与前阶段相同。

可能有某些 α_j（在第 j 个方向移动的净距离）等于 0，在这种场合上述正交化过程必须修改[9]。一个不同的而且更经济的正交化过程已由 Palmer[9] 提出。当 $\|u^1\|<\varepsilon$ 对连续几个勘探阶段都成立时，这个算法就可结束，这里 ε 是某个小的正数。

例 9.3.1

考虑用 Rosenbrock 旋转方向法求下列函数极小值点的问题：

$$f(x_1, x_2)=(x_1)^2+(x_2)^2-3x_1-x_1x_2+3. \tag{9.24}$$

假定我们从 $x_0^1=(0, 0)$ 出发，并且首先沿坐标轴进行勘探移动，即

$$\zeta^1=\begin{pmatrix}1\\0\end{pmatrix}, \qquad \zeta^2=\begin{pmatrix}0\\1\end{pmatrix}, \tag{9.25}$$

还假设 $\alpha=3$，$\beta=0.5$。勘探移动的位置和结果汇总于表 9.1 中。

表 9.1 第一个勘探阶段

方 向	经过的点	函数值	结 果
—	(0.0, 0.0)	3.0	—
ζ^1	(1.0, 0.0)	1.0	成 功
ζ^2	(1.0, 1.0)	1.0	成 功
ζ^1	(4.0, 1.0)	4.0	失 败
ζ^2	(1.0, 4.0)	13.0	失 败

在两个搜索方向都获得一次成功并接着一次失败之后，我们建立一个新的出发点 $x_B^1(1,\ 1)$及旋转方向. 注意到 $\alpha_1=1, \alpha_2=1$, 利用(9.17)到(9.23), 得到

$$u^1=1\begin{pmatrix}1\\0\end{pmatrix}+1\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}, \tag{9.26}$$

$$u^2=\qquad 1\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix}, \tag{9.27}$$

$$w^1=\begin{pmatrix}1\\1\end{pmatrix}, \qquad \hat{\zeta}^1=\begin{pmatrix}\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}\end{pmatrix}\cong\begin{pmatrix}0.7071\\0.7071\end{pmatrix}, \tag{9.28}$$

$$w^2=\begin{pmatrix}0\\1\end{pmatrix}-\left[(0.1)\begin{pmatrix}\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}\end{pmatrix}\right]\begin{pmatrix}\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}\end{pmatrix}=\begin{pmatrix}-\frac{1}{2}\\ \frac{1}{2}\end{pmatrix}, \tag{9.29}$$

$$\hat{\zeta}^2=\begin{pmatrix}-\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}\end{pmatrix}\cong\begin{pmatrix}-0.7071\\0.7071\end{pmatrix}. \tag{9.30}$$

从 x_B^1 开始一个新的勘探阶段，移动及其结果列于表 9.2 中.

表 9.2 第二个勘探阶段

方向	经 过 的 点	函 数 值	结 果
—	(1.0000, 1.0000)	1.0000	—
ξ^1	(1.7071, 1.7071)	0.7929	成 功
ξ^2	(1.0000, 2.4142)	4.4142	失 败
ξ^1	(3.8284, 3.8284)	6.1716	失 败
ξ^2	(2.0606, 1.3536)	0.1072	成 功
ξ^1	(1.0000, 0.2929)	0.7929	失 败
ξ^2	(3.1213, 0.2929)	2.5503	失 败

在第二个勘探阶段中得到的最好点是 $x_B^2=(2.0606, 1.3536)$，一个新的勘探阶段将在旋转了搜索方向之后从此点出发. f 的真正的极小值点在 $x^*=(2.0, 1.0)$，此点的函数值是 $f(x^*)=0$. ∎

Davies、Swann 和 Campey[16] 修改了 Rosenbrock 方法，把上述勘探阶段改为对这 n 个正交方向依次进行一维搜索. 搜索方向的旋转还如 Rosenbrock 方法一样进行. 对其中的一维搜索，[2] 中推荐用一种单个的二次逼近法，这种方法在上章中曾叙述过.

这个 Davies、Swann 和 Campey 方法具有一个有趣的性质，这个性质的一般化在随后将要讨论的几乎所有无约束极小化方法中都具有基本的重要性. 假定被极小化的函数是二次的，并由下式给出：

$$f(x)=a+b^Tx+\frac{1}{2}x^TQx, \quad x\in R^n, \tag{9.31}$$

其中 Q 是一个对称正定阵. 那么 f 有一个整体极小值点，为

$$x^*=-Q^{-1}b. \tag{9.32}$$

如果 $Q=cI$，其 $c>0$ 是一个实数而 I 是一个单位阵，则 $f(x)=$ 常数所表示的等值面是一个中心在 x^* 的超球面. 从 R^n 中任意一点出发，沿着任何 n 个相互正交的方向依次进行一维搜索，我们就可以精确地确定 f 的极小值点. 我们用一个例子来说明这个性质.

例 9.3.2

设二次函数 f 由下式给出：

$$f(x)=3(x_1)^2+3(x_2)^2-3x_1+4x_2. \tag{9.33}$$

这个函数的整体极小值点在 $x^*=\left(\frac{1}{2},\ -\frac{2}{3}\right)$. 假设我们从 $x_B^0=(-3,\ 1)$ 出发. 两个正交方向是

$$\zeta^1=\begin{pmatrix} \dfrac{2}{\sqrt{5}} \\ -\dfrac{1}{\sqrt{5}} \end{pmatrix},\quad \zeta^2=\begin{pmatrix} \dfrac{1}{\sqrt{5}} \\ \dfrac{2}{\sqrt{5}} \end{pmatrix}. \tag{9.34}$$

首先找 f 在 ζ^1 方向的极小值点:

$$f(x_B^0+\theta_1\zeta^1)=3\left(-3+\frac{2\theta_1}{\sqrt{5}}\right)^2+3\left(1-\frac{\theta_1}{\sqrt{5}}\right)^2 -3\left(-3+\frac{2\theta_1}{\sqrt{5}}\right)+4\left(1-\frac{\theta_1}{\sqrt{5}}\right). \tag{9.35}$$

因为这是关于变量 θ_1 的一个二次函数,所以二阶多项式逼近将是精确的(忽略舍入误差). (9.35)的极小值在

$$\theta_1^*=\frac{26\sqrt{5}}{15} \tag{9.36}$$

达到,从而在 ζ^1 方向上的最好点是 $t_1^1=(t_{11}^1,\ t_{12}^1)$,其中

$$t_{11}^1=-3+\frac{(2)(26)}{15}=\frac{7}{15}, \tag{9.37}$$

$$t_{12}^1=1-\frac{26}{15}=-\frac{11}{15}. \tag{9.38}$$

从这点出发沿 ζ^2 方向进行,再对下列函数进行一维搜索以求其极小值点:

$$f(t_1^1+\theta_2\zeta^2)=3\left(\frac{7}{15}+\frac{\theta_2}{\sqrt{5}}\right)^2+3\left(-\frac{11}{15}+\frac{2\theta_2}{\sqrt{5}}\right)^2 -3\left(\frac{7}{15}+\frac{\theta_2}{\sqrt{5}}\right)+4\left(-\frac{11}{15}+\frac{2\theta_2}{\sqrt{5}}\right). \tag{9.39}$$

这也是 θ_2 的一个二次函数,(9.39)的极小值在

$$\theta_2^*=\frac{\sqrt{5}}{30} \tag{9.40}$$

达到. 因此

$$t_{21}^1=\frac{7}{15}+\frac{1}{30}=\frac{1}{2}, \tag{9.41}$$

$$t_{22}^1=-\frac{11}{15}+\frac{2}{30}=-\frac{2}{3}. \tag{9.42}$$

正如所断言的，点 $t_2^1=(t_{21}^1,\ t_{22}^1)$ 是 f 的无约束极小值点. ▌

对于一般带有正定阵 Q 的二次函数，单是相继沿着 n 个相互正交的方向进行一轮搜索，一般是无法确定它的精确的极小值点的. 正如下节将要指出的，共轭方向——它可以看作是推广了的正交方向——具有下列性质：用一种与上述类似的方法，对于 R^n 上定义的一个二次函数可以通过至多 n 次一维搜索而达到它的精确的极小值点. 在本章的以后部分我们将介绍基于共轭方向的一类极小化方法中的第一个，对于 R^n 上的一般函数来说，它已成为不用计算导数的最好的极小化方法之一.

9.4 共 轭 方 向

我们现在对二次函数引进一个重要的概念，它是整整一类无约束极小化方法的基础.

两个向量 $x\in R^n$, $y\in R^n$ 称为是关于 $n\times n$ 对称正定阵 A 的**共轭方向**，如果

$$x^TAy=0. \tag{9.43}$$

共轭方向的概念是正交性（当 A 是 $n\times n$ 单位阵的场合）的一个推广. 众所周知，一个对称的 $n\times n$ 阵 A 具有 n 个正交的特征向量[8]. 这 n 个向量也是相互共轭的，因为若 x^1, x^2 是 A 的这种特征向量，那么 $Ax^2=\lambda x^2$，这里 λ 是相应的特征值，因而

$$(x^1)^TAx^2=(x^1)^T\lambda x^2=\lambda(x^1)^T(x^2)=0. \tag{9.44}$$

因此对每个 $n\times n$ 对称正定阵，至少有 n 个相互共轭的方向.

我们可以从 R^n 中 n 个线性独立的向量 v^1, v^2, $\cdots$, v^n 出发，用一种与以前介绍的 Gram-Schmidt 正交化方法相类似的过程来构造关于 A 的相互共轭的 n 个方向.

令

$$z^1 = u^1, \tag{9.45}$$

$$z^j = u^j - \sum_{k=1}^{j-1} \frac{(u^j)^T A z^k}{(z^k)^T A z^k} z^k, \quad j = 2, \cdots, n, \tag{9.46}$$

则向量 $z^1, \cdots, z^n$ 是关于 A 相互共轭的. 希望读者自己证明, 上述关系式的确产生 n 个共轭方向, 并且关于 A 相互共轭的每一组非零向量 $z^1, \cdots, z^k$ 是线性独立的. 应当指出, 共轭方向的定义也可以不用正定性, 不过一些最有意义的结果都依赖于这个性质.

共轭向量的一个几何解释如下: 设 f 是由(9.31)给定的二次函数, 并设 Q 是一个正定阵. 设 x^* 是对于所有 $x \in R^n$ 使 $f(x)$ 达到极小值的点, 那么曲面 $f(x) = c$(常数)通常是以 x^* 为中心的椭球面. 若 x^0 是满足 $f(x^0) = c$ 的一个点, 在 x^0 构造一个切于 $f(x) = c$ 的超平面, 则连结 x^0 与 x^* 的向量关于 Q 共轭于切平面上的每一个向量. 二维的情形说明于图 9.4.

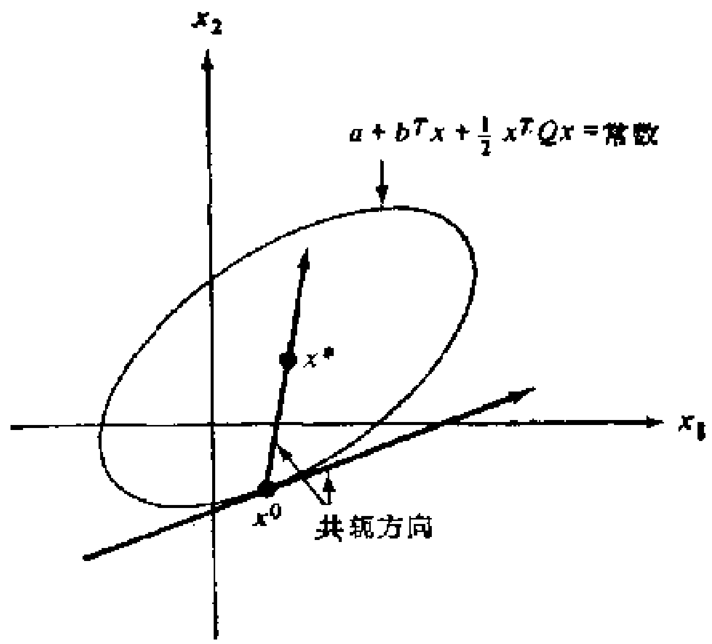


图 9.4 关于 Q 的共轭方向

下面假定在 R^n 中有 m 个非零向量 $z^1, \cdots, z^m$, $m \leqslant n$, 它们关于正定阵 Q 相互共轭. 那么, 这些向量是线性独立的, 而且满足下列关系式的向量的全体构成 R^n 的一个 m 维子空间:

$$x=\sum_{j=1}^{m}\alpha_j z^j, \tag{9.47}$$

其中 α_j 是任意实数. 对于一个给定的点 $x^0\in R^n$, 满足

$$x=x^0+\sum_{j=1}^{m}\alpha_j z^j, \tag{9.48}$$

(其中 z^j 是 m 个线性独立向量而 α_j 是任意数)的向量集合是一个仿射集或线性流形(在第 4 章也定义过), 称它是由点 x^0 及向量 z^1, …, z^m 所生成的. 当 $m=n$ 时, 仿射集变成整个空间 R^n.

现在我们可以叙述和证明一个把共轭方向和二次函数无约束极小化联系起来的重要结果, 这个结果属于 Powell[11].

定理 9.1

若 z^1, …, z^m 是非零的、关于正定阵 Q 相互共轭的方向, 则二次函数

$$f(x)=a+b^Tx+\frac{1}{2}x^TQx,\quad x\in R^n \tag{9.49}$$

在由点 $x^0\in R^n$ 及向量 z^1, …, z^m 生成的仿射集上的极小值点可以通过沿每个共轭方向搜索一次而找到.

【证明】 这个仿射集中的每一个点 x 由(9.48)给出, 所求的极小值点是点 $x^0+\alpha_1^*z^1+\cdots+\alpha_m^*z^m$, 其中 α_j 选得使下列函数达到极小:

$$\begin{aligned}f(x)&=f\left(x^0+\sum_{j=1}^{m}\alpha_j z^j\right)\\&=f(x^0)+\sum_{j=1}^{m}\left[\alpha_j(z^j)^T(Qx^0+b)+\frac{1}{2}(\alpha_j)^2(z^j)^TQz^j\right].\end{aligned} \tag{9.50}$$

注意到非零共轭方向是线性独立的, 因此 z^1, …, z^m 生成一个 m 维子空间. 因为没有 $\alpha_j\alpha_k$ 项($j\neq k$), 最优的 α_j 通过下列 m 个极小化问题而找到:

$$\min_{\alpha_j}\left\{f(x^0)+\alpha_j(z^j)^T(Qx^0+b)+\frac{1}{2}(\alpha_j)^2(z^j)^TQz^j\right\},\quad j=1,\ \cdots,\ m. \tag{9.51}$$

但是这个问题等价于求解

$$\min_{\alpha_j} f(x^0+\alpha_j z^j), \quad j=1, \cdots, m. \tag{9.52}$$

因此，沿 m 个方向 $z^1, \cdots, z^m$ 各搜索一次极小值点将产生最优的 α_j^*，因而得到 f 在该仿射集上的极小值点. ▌

若 $m=n$，由定理 9.1 推出二次函数(9.49)在 R^n 上的极小值点可通过沿 n 个非零的、关于 Q 相互共轭的方向搜索而得到.

例 **9.4.1**

我们用一个双变量的二次函数来证实上述结论. 设

$$f(x)=2(x_1)^2+6(x_2)^2+2x_1x_2+2x_1+3x_2+3. \tag{9.53}$$

这函数可改写成(9.49)，其中 $a=3$, $b=(2, 3)^T$，而

$$Q=\begin{bmatrix} 4 & 2 \\ 2 & 12 \end{bmatrix}. \tag{9.54}$$

假定选 $(z^1)^T=(1,0)$，关于 Q 共轭于 z^1 的方向是

$$(z^2)^T=\left(-\frac{1}{2}, 1\right).$$

我们来找 f 在由点 $x^0=(0, 0)$ 及向量 z^1, z^2 生成的仿射集上的极小值点，在当前的场合，这个仿射集显然是整个空间 R^2. 按定理 9.1，必须在 z^1 及 z^2 方向进行两个独立的一维搜索. 我们可以用练习 8.D 所提供的方法来进行这些搜索，因为在这种场合下二次逼近将是精确的. 从 z^1 方向开始，我们要极小化 $F_1(\theta^1)=f(x^0+\theta^1 z^1)$. 可以取 $\theta_1^1=0$, $\theta_2^1=1$, $\theta_3^1=2$，而 $F_1(0)=3$, $F_1(1)=7$, $F_1(2)=15$. 因此

$$\theta^{1*}=\frac{(-3)(3)+(4)(7)+(-1)(15)}{2[(-1)(3)+(2)(7)+(-1)(15)]}=-\frac{1}{2}. \tag{9.55}$$

现在沿 z^2 方向进行，我们来极小化 $F_2(\theta^2)=f(x^0+\theta^2 z^2)$. 再一次取 $\theta_1^2=0$, $\theta_2^2=1$, $\theta_3^2=2$. 而 $F_2(0)=3$, $F_2(1)=\frac{21}{2}$, $F_2(2)=29$，因此

$$\theta^{2*}=\frac{(-3)(3)+(4)(21/2)+(-1)(29)}{2[(-1)(3)+(2)(21/2)+(-1)(29)]}=-\frac{2}{11}. \tag{9.56}$$

这样一来，f 在 R^2 上的极小值点给出如下：

$$x^* = x^0 + \theta^{1*}z^1 + \theta^{2*}z^2$$

$$= \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -\frac{1}{2} \\ 0 \end{pmatrix} + \begin{pmatrix} \frac{1}{11} \\ -\frac{2}{11} \end{pmatrix} = \begin{pmatrix} -\frac{9}{22} \\ -\frac{2}{11} \end{pmatrix}. \qquad (9.57)$$

读者可以通过对 f 进行微分来验证这个结果. ∎

两个仿射集 S 与 T, $S \neq T$, 称为是平行的, 如果它们由相同的方向 $z^1, \cdots, z^m$ 与不同的点 $x(S) \in S$ 及 $x(T) \in T$ 所生成. 我们有下述定理.

定理 9.2

若 $x^*(S)$ 与 $x^*(T)$ 分别是两个平行的仿射集 S 与 T 上使函数 $f(x)$(它由(9.49)给出)达到极小值的点, 则向量 $x^*(T) - x^*(S)$ 关于 Q 共轭于 S 和 T 中的任何一个向量.

【证明】 设 z 是在 S 和 T 中的一个向量, 则对 $\alpha = 0$ 成立

$$\frac{d}{d\alpha}[f(x^*(S) + \alpha z)] = 0, \qquad (9.58)$$

或等价地为

$$z^T[Qx^*(S) + b] = 0. \qquad (9.59)$$

类似地有

$$z^T[Qx^*(T) + b] = 0. \qquad (9.60)$$

从(9.60)减去(9.59)得到

$$z^T Q[x^*(T) - x^*(S)] = 0. \qquad (9.61)$$

∎

这两个定理是证明下节所述的 Powell 法所必需的, Powell 法具有以后将要定义的二次终结性质.

9.5 Powell 法

前面介绍的三种算法都是为了寻找一般无约束函数的极小值点而导出的, 通常需要无穷次迭代. 另一方面, 如果我们有一种方法, 它可以用有限步找到 R^n 上二次函数的极小值点, 那么这种方

法对于极小化一个具有连续二阶导数的一般函数也将是有效的.在本章的其余部分和下一章中我们提供一整类方法，这类方法都具有二次终结的性质：用这种方法来找 R^n 上由(9.31)给定的具有正定阵 Q 的二次函数的极小值点时，至多在 n 步内终结. 这个由一个算法规定的"步"将广义地解释为从 R^n 中的一个点移动到另一个点，显然在许多情形，这些移动只有通过一个无穷的过程才能精确地进行. 譬如，当一个算法由若干个沿 R^n 中某个方向极小化一个函数所确定的步所组成时，便是这种情形. 因此，例如用黄金分割法来沿着一个方向求一个二次函数的极小，就要作无穷次函数值计算. 另一方面，如用第 8 章所述的多项式逼近法(忽略舍入误差)，便可通过有限次算术运算得到这一二次函数的精确的极小值点. 然而，一般地讲，一维极小化是一个无穷的过程.

开始时，Powell 法作为一个经验方法来介绍，正如本章前些方法一样. 它的基于共轭方向的基本原则将在本节稍后加以阐明.

Powell 法[11]的第一个和基本的做法可叙述如下：过程的每一阶段由 $n+1$ 个相继的一维直线搜索构成，先沿着 n 个线性独立的方向，然后沿着由这一阶段的出发点与最好点(在 n 次一维线性搜索后得到的)相连结的方向. 在这些搜索之后，前 n 个方向之一被代之以这第 $n+1$ 个方向，然后开始一个新的阶段.

这方法的第 k 阶段用下列步骤给出. 设 $x_B^{k-1}=t_0^k\in R^n$ 是第 k 阶段的出发点，并设给定了 n 个线性独立的方向 $\Delta_1^k, \cdots, \Delta_n^k$ (对 $k=1$，通常选坐标方向). 对 $j=1, \cdots, n$，找数 θ_j^* 使

$$f(t_{j-1}^k+\theta_j^*\Delta_j^k)=\min_{\theta_j} f(t_{j-1}^k+\theta_j\Delta_j^k). \tag{9.62}$$

令

$$t_j^k=t_{j-1}^k+\theta_j^*\Delta_j^k, \quad j=1, \cdots, n, \tag{9.63}$$

$$\Delta_j^{k+1}=\Delta_{j+1}^k, \quad j=1, \cdots, n-1, \tag{9.64}$$

$$\Delta_n^{k+1}=\Delta_{n+1}^k=t_n^k-t_0^k. \tag{9.65}$$

找 θ_{n+1}^* 使

$$f(t_n^k+\theta_{n+1}^*(t_n^k-t_0^k))=\min_{\theta_{n+1}} f(t_n^k+\theta_{n+1}(t_n^k-t_0^k)), \qquad (9.66)$$

而且令

$$x_B^k=t_n^k+\theta_{n+1}^*(t_n^k-t_0^k), \qquad (9.67)$$

若 $\|x_B^k-x_B^{k-1}\|<\varepsilon$（这里 $\varepsilon>0$ 是某个预先给定的数），则停止；否则，进入第 $k+1$ 阶段.

这个算法的少数几步在图 9.5 中说明，其中求极小的是一个双变量的一般函数.

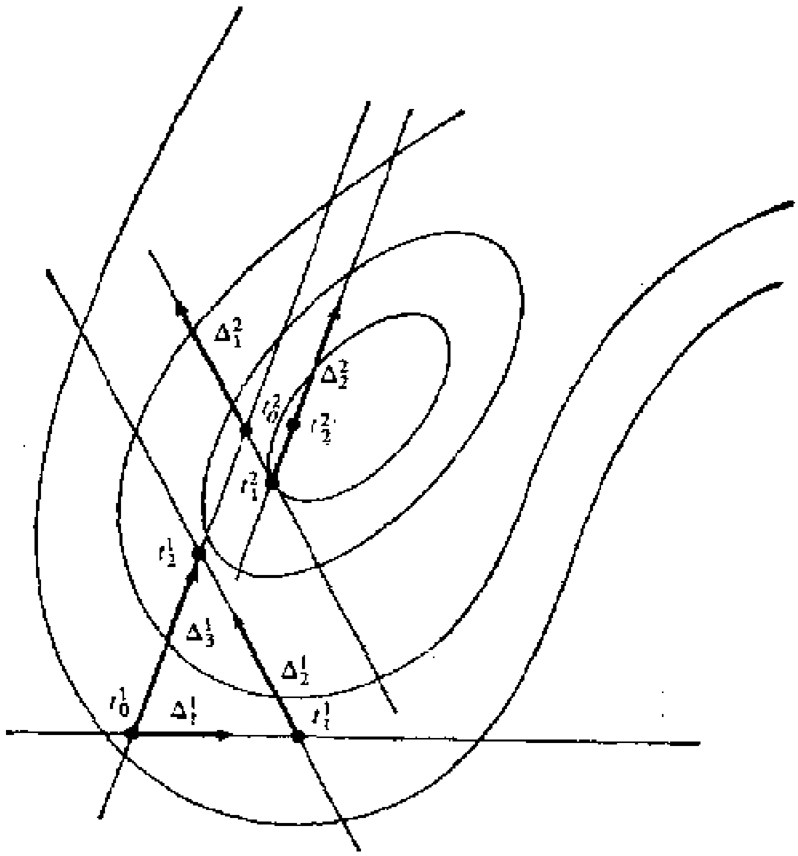


图 9.5 Powell 法

以后将证明 Powell 算法可在至多 n 个阶段中找到 R^n 上具有正定阵 Q 的二次函数的精确极小值点，只要对于 $k=1, \cdots, n$，这些搜索方向 $\Delta_1^k, \cdots, \Delta_n^k$ 都是线性独立的.

例 9.5.1

通过极小化下列双变量二次函数：

$$f(x)=\frac{3}{2}(x_1)^2+\frac{1}{2}(x_2)^2-x_1x_2-2x_1, \qquad (9.68)$$

来说明 Powell 法. 这函数的极小值点在 $x^*=(1, 1)$. 假定从 $x_B^0=t_0^1=(-2, 4)$出发沿着坐标方向进行搜索:

$$\Delta_1^1=\begin{pmatrix}1\\0\end{pmatrix}, \quad \Delta_2^1=\begin{pmatrix}0\\1\end{pmatrix}, \tag{9.69}$$

第一个极小化是沿 Δ_1^1 方向:

$$\begin{aligned}\min_{\theta_1} f(t_0^1+\theta_1\Delta_1^1) &= \frac{3}{2}(-2+\theta_1)^2+\frac{1}{2}(4)^2\\ &\quad -4(-2+\theta_1)-2(-2+\theta_1),\end{aligned} \tag{9.70}$$

在 $\theta_1^*=4$ 达到最优, 因此 $t_1^1=(2, 4)$. 现在, 沿 Δ_2^1 方向极小化:

$$\min_{\theta_2} f(t_1^1+\theta_2\Delta_2^1) = \frac{3}{2}(2)^2+\frac{1}{2}(4+\theta_2)^2-2(4+\theta_2)-4, \tag{9.71}$$

得到 $\theta_2^*=-2$. 因此 $t_2^1=(2, 2)$, 从而

$$\Delta_3^1=\begin{pmatrix}2-(-2)\\2-4\end{pmatrix}=\begin{pmatrix}4\\-2\end{pmatrix}, \tag{9.72}$$

我们再沿这个方向进行极小化:

$$\begin{aligned}\min_{\theta_3} f(t_2^1+\theta_3\Delta_3^1) &= \frac{3}{2}(2+4\theta_3)^2+\frac{1}{2}(2-2\theta_3)^2\\ &\quad -(2+4\theta_3)(2-2\theta_3)-2(2+4\theta_3).\end{aligned} \tag{9.73}$$

最优点是在 $\theta_3^*=-\frac{2}{17}$, 因而

$$x_B^1=t_0^2=\begin{pmatrix}2-\frac{8}{17}\\[2ex] 2+\frac{4}{17}\end{pmatrix}=\begin{pmatrix}\frac{26}{17}\\[2ex] \frac{38}{17}\end{pmatrix}. \tag{9.74}$$

算法的第一轮迭代至此结束. 第二轮迭代的前两个搜索方向按(9.64)及(9.65)应为

$$\Delta_1^2=\begin{pmatrix}0\\1\end{pmatrix}, \quad \Delta_2^2=\begin{pmatrix}4\\-2\end{pmatrix}. \tag{9.75}$$

我们求解

$$\min_{\theta_1} f(t_0^2+\theta_1\Delta_1^2)=\frac{3}{2}\left(\frac{26}{17}\right)+\frac{1}{2}\left(\frac{38}{17}+\theta_1\right)^2$$

$$-\frac{26}{17}\left(\frac{38}{17}+\theta_1\right)-\frac{52}{17}, \tag{9.76}$$

找到 $\theta_1^*=-\frac{12}{17}$, $t_1^2=\left(\frac{26}{17}, \frac{26}{17}\right)$. 其次, 求解

$$\min_{\theta_2} f(t_1^2+\theta_2\Delta_2^2)=\frac{3}{2}\left(\frac{26}{17}+4\theta_2\right)^2+\frac{1}{2}\left(\frac{26}{17}-2\theta_2\right)^2$$

$$-\left(\frac{26}{17}+4\theta_2\right)\left(\frac{26}{17}-2\theta_2\right)-2\left(\frac{26}{17}+4\theta_2\right), \tag{9.77}$$

最优点在 $\theta_2^*=-\frac{18}{289}$, 因此, $t_2^2=\left(\frac{370}{289}, \frac{478}{289}\right)$.

现在, 我们找出第二轮迭代的第三个方向

$$\Delta_3^2=\begin{pmatrix}\frac{370}{289}-\frac{26}{17}\\[2mm] \frac{478}{289}-\frac{38}{17}\end{pmatrix}=-\begin{pmatrix}\frac{72}{289}\\[2mm] \frac{168}{289}\end{pmatrix}, \tag{9.78}$$

并求解

$$\min_{\theta_3} f(t_2^2+\theta_3\Delta_3^2)=\frac{3}{2}\left(\frac{370}{289}-\frac{72}{289}\theta_3\right)^2+\frac{1}{2}\left(\frac{478}{289}-\frac{168}{289}\theta_3\right)^2$$

$$-\left(\frac{370}{289}-\frac{72}{289}\theta_3\right)\left(\frac{478}{289}-\frac{168}{289}\theta_3\right)$$

$$-2\left(\frac{370}{289}-\frac{72}{289}\theta_3\right), \tag{9.79}$$

最优解在 $\theta_3^*=\frac{9}{8}$, 因而 $x_3^2=(1, 1)$. 因此, 正如所断言的, 二次函数的精确极小值点在两轮迭代中找到. ❚

注意, 这个例子中的搜索方向 Δ_1^k 和 $\Delta_2^k(k=1, 2)$ 是线性独立的. 这个条件十分重要, 因为有例子可以说明: 如上所述的 Powell 方法的第一个做法可以经 n 轮迭代仍未到达 f 的极小值点, 甚至可能经任意次迭代还是达不到它.

例 9.5.2 [17]

考察 R^3 上如下定义的函数 f:

$$f(x_1, x_2, x_3) = (x_1 - x_2 + x_3)^2 + (-x_1 + x_2 + x_3)^2 + (x_1 + x_2 - x_3)^2. \tag{9.80}$$

这是一个严格凸的二次函数，具有一个唯一的极小值点 $x^* = (0, 0, 0)$. 取 $x_B^0 = \left(\frac{1}{2}, 1, \frac{1}{2}\right)$, 并用下列方向开始施行 Powell 方法,

$$\Delta_1^1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \Delta_2^1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \Delta_3^1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{9.81}$$

前三步及其结果汇总在表 9.3 中:

表 9.3 用 Powell 方法的前三步

j	方向 Δ_j^1	t_j^1	函数值 $f(t_j^1)$
0	—	$\left(\frac{1}{2}, 1, \frac{1}{2}\right)$	2
1	Δ_1^1	$\left(\frac{1}{2}, 1, \frac{1}{2}\right)$	2
2	Δ_2^1	$\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{2}\right)$	$\frac{2}{3}$
3	Δ_3^1	$\left(\frac{1}{2}, \frac{1}{3}, \frac{5}{18}\right)$	$\frac{42}{81}$

新的方向将是

$$t_3^1 - t_0^1 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{3} \\ \frac{5}{18} \end{pmatrix} - \begin{pmatrix} \frac{1}{2} \\ 1 \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{2}{3} \\ -\frac{2}{9} \end{pmatrix}, \tag{9.82}$$

因此得到

$$\Delta_1^2=\begin{pmatrix}0\\1\\0\end{pmatrix},\quad \Delta_2^2=\begin{pmatrix}0\\0\\1\end{pmatrix},\quad \Delta_3^2=\begin{pmatrix}0\\-\frac{2}{3}\\-\frac{2}{9}\end{pmatrix}. \tag{9.83}$$

注意,所有这三个新的搜索方向的第一个分量都等于0,因此所有即将到来的点的第一个分量都将保持为 $\frac{1}{2}$,这样就永远不能达到真正的最优点 $x^*=(0, 0, 0)$,其原因是,向量 $\Delta_j^2(j=1, 2, 3)$是线性相关的,因而不能张成整个空间 R^3.]

读者可以试着去构造这种类型的其他例子,只要使搜索的出发点已是函数沿第一个搜索方向的极小值点即可.

现在我们来说明,怎样可用定理9.1与定理9.2(它们涉及共轭方向的性质)来证明 Powell 法在适当假定下是二次终结的. 假设已经给定了 R^n 上一个带有正定阵 Q 的二次函数 f,一个出发点 $x_B^0\in R^n$,以及 n 个线性独立的方向 $\Delta_1^1, \cdots, \Delta_n^1$. 在进行了第一阶段的步骤之后,我们有了一个方向 $z^1=t_n^1-t_0^1=\Delta_n^2$ 及第二阶段的新的出发点 $x_B^1=t_0^2$. 注意,如果 R^n 中的一个点在 n 个线性独立方向上都是最优的,那末它必是该二次函数的整体最优点. 设 $t_n^1\neq t_0^1$;即方向 z^1 不是零. 由(9.67)知,点 x_B^1 是 z^1 方向的一个极小值点;再设方向 $\Delta_1^2, \cdots, \Delta_n^2$ 是线性独立的,就可以达到一点 t_n^2,它也是沿 z^1 方向的一个极小值点,而且包含在一个平行的仿射集中. 根据定理9.2,方向 $z^2=t_n^2-t_0^2$ 是关于 Q 共轭于 z^1 的. 现在假定算法的 k 个阶段已完成并产生了 k 个非零的、关于 Q 相互共轭的方向 z^1, $z^2, \cdots, z^k$. 若方向 $\Delta_1^k, \cdots, \Delta_{n-k}^k, z^1, \cdots, z^k$ 是线性独立的,那么根据定理9.1及定理9.2,方向 $z^{k+1}=t_n^{k+1}-t_0^{k+1}$ 与 $z^1, \cdots, z^k$ 是相互共轭的. 在完成了 n 个这样的阶段之后,所有的搜索方向是关于 Q 相互共轭的,因此根据定理9.1,达到了 f 在 R^n 上的极小值点.

9.6 避免线性相关的搜索方向

Powell 注意到,即使在不如例9.5.2那样极端的情况下,他的

方法也可能选取到近乎线性相关的搜索方向，特别是有大量变量时更是如此，这种可能性对收敛性来说将有严重的后果。为了避免这个困难，他提出了这个方法的一种修改。这个修改了的做法不再具有二次终结性质，但它的性能一般却是非常满意的.

修改的 Powell 方法[11]（又为 Sargent[14] 稍加改善）的第 k 个阶段用下述规则给定. 再一次用 $x_B^{k-1}=t_0^k$ 表示第 k 阶段的出发点，并设给定了 n 个线性独立的方向 $\Delta_1^k, \cdots, \Delta_n^k$. 首先，正如第一种做法那样，对 $j=1, \cdots, n$ 找 t_j^k. 其次，找指标 m 使得

$$f(t_{m-1}^k)-f(t_m^k)=\max_{j=1,\cdots,n}\{f(t_{j-1}^k)-f(t_j^k)\}. \tag{9.84}$$

令 $\Delta_{n+1}^k=(t_n^k-t_0^k)$；若 $\|t_n^k-t_0^k\|\leqslant\varepsilon$，就停止；否则找一个数 α_{n+1}^* 使得

$$f(t_0^k+\alpha_{n+1}^*\Delta_{n+1}^k)=\min_{\alpha_{n+1}} f(t_0^k+\alpha_{n+1}\Delta_{n+1}^k), \tag{9.85}$$

并且记 $t_0^{k+1}=x_B^k=t_0^k+\alpha_{n+1}^*\Delta_{n+1}^k$. 若 $\|x_B^k-x_B^{k-1}\|\leqslant\varepsilon$，就停止，并认为已收敛到一个极小值点；否则，若

$$|\alpha_{n+1}^*|<\left[\frac{f(t_0^k)-f(t_0^{k+1})}{f(t_{m-1}^k)-f(t_m^k)}\right]^{1/2}, \tag{9.86}$$

则取

$$\Delta_j^{k+1}=\Delta_j^k, \quad j=1, \cdots, n. \tag{9.87}$$

换句话说，第 $k+1$ 阶段的搜索方向保持与第 k 阶段相同. 若 (9.86) 不成立，则取

$$\Delta_j^{k+1}=\Delta_j^k, \qquad j=1, \cdots, m-1, \tag{9.88}$$

$$\Delta_j^{k+1}=\Delta_{j+1}^k, \quad j=m, \cdots, n, \tag{9.89}$$

并进入第 $k+1$ 阶段.

例 9.6.1

我们用上例的二次函数来说明修改了的 Powell 法. 第一步与那里相同. 可以看到从 t_1^1 到 t_2^1 时有最大的函数值减少；因此 $m=2$，我们有

$$\Delta_4^1=\begin{pmatrix}0\\ -\dfrac{2}{3}\\ -\dfrac{2}{9}\end{pmatrix}. \tag{9.90}$$

接着找一个α_4^*使其极小化$f\left(\frac{1}{2},\ 1-\frac{2}{3}\alpha_4,\ \frac{1}{2}-\frac{2}{9}\alpha_4\right)$，容易算得$\alpha_4^*=\frac{9}{8}$，因此

$$t_0^2=\begin{pmatrix}\frac{1}{2}\\1\\\frac{1}{2}\end{pmatrix}+\frac{9}{8}\begin{pmatrix}0\\-\frac{2}{3}\\-\frac{2}{9}\end{pmatrix}=\begin{pmatrix}\frac{1}{2}\\\frac{1}{4}\\\frac{1}{4}\end{pmatrix}, \tag{9.91}$$

而$f(t_0^2)=\frac{1}{2}$. 现在

$$\left[\frac{f(t_0^1)-f(t_0^2)}{f(t_1^1)-f(t_2^1)}\right]^{1/2}=\left(\frac{2-\frac{1}{2}}{2-\frac{2}{3}}\right)^{1/2}=\left(\frac{9}{8}\right)^{1/2}. \tag{9.92}$$

因为

$$\alpha_4^*>\left(\frac{9}{8}\right)^{1/2}, \tag{9.93}$$

可以看到(9.86)并不成立. 相应地，新的方向将是

$$\Delta_1^2=\begin{pmatrix}1\\0\\0\end{pmatrix},\quad \Delta_2^2=\begin{pmatrix}0\\0\\1\end{pmatrix},\quad \Delta_3^2=\begin{pmatrix}0\\-\frac{2}{3}\\-\frac{2}{9}\end{pmatrix}, \tag{9.94}$$

而且这些向量是线性独立的. 下一阶段的计算留给读者. ▮

我们现在来导出某些结果，通过这些结果可以阐明修改的 Powell 方法. 我们已经看到，沿着共轭方向来搜索二次函数的极小值点是我们所希望的. 首先，我们要指明一组向量关于一个正定对称阵接近于相互共轭的程度是如何度量的. 其次，我们将会看到，对于要寻找极小值点的二次函数的 Hesse 阵 Q 来说，用修改的 Powell 方法产生的搜索方向是趋向于相互共轭的，而且在每次迭代中这些方向是线性独立的. 在指出了这个方法对二次函数的

有效性之后，我们将对一大类函数证明一个简化了的做法的收敛性。

先回忆一下正交阵的定义及其某些性质。

一个实矩阵 P 如果适合 $P^TP=PP^T=I$（单位矩阵），则称它为正交阵[8]。正交阵的某些基本性质是 $P^{-1}=P^T$ 和 $|\det P|=1$. P 的每行与每列都是单位长度（规范化了）的向量。进一步说，P 是正交的充要条件为 P 的行与列是相互规范正交的。给定任意向量 x，令 $y=Px$，则 $\|y\|=\|x\|$. 设 x 与 y 是任意向量，令 $x^1=Px$，$y^1=Py$，则 $(x^1)^Ty^1=x^TP^TPy=x^Ty$. 因此在正交变换下数量积是不变的。特别地，通过矩阵 P，正交向量将变换为正交向量。对于正交阵与正交变换的其他性质读者可参考线性代数的标准教科书。下面是 Powell[11] 的一个结果，它用一个行列式的值来给出一组搜索方向接近于相互共轭的程度的一个适当的度量。这个结果既构成了 Powell 方法的修改了的做法（正如上述）的基础，也构成了本节稍后将概述的一个新算法的基础。

定理 9.3

设 G 是一个实的 $n\times n$ 对称正定阵，而 $d^1, \cdots, d^n$ 是 n 个向量，它们按下式意义是规范化的：

$$(d^i)^TGd^i=1, \quad i=1, \cdots, n. \tag{9.95}$$

设 D 是以向量 d^i 为列作出的矩阵，则当且仅当向量 d^i 关于 G 为相互共轭时，$|\det D|$ 达到它的最大值。

【证明】 我们先证明：若(9.95)成立，而且 $|\det D|$ 达到最大，则 d^i 必定关于 G 共轭。设 d^1 与 d^2 是规范化的向量，它们关于 G 不共轭，也即

$$(d^1)^TGd^2=c\neq 0. \tag{9.96}$$

因为 G 是正定的，它有一个“平方根”矩阵 B；也就是说，$G=B^TB$. 用 Cauchy-Schwarz 不等式[8]，

$$\begin{aligned}|(d^1)^TGd^2| &= |(d^1)^TB^TBd^2| \\ &\leqslant [(Bd^1)^TBd^1]^{1/2}[(Bd^2)^TBd^2]^{1/2} \qquad (9.97)\\ &= [(d^1)^TGd^1]^{1/2}[(d^2)^TGd^2]^{1/2}=1. \qquad (9.98)\end{aligned}$$

因此$|c|\leqslant 1$，等号成立的充要条件为

$$d^1=\frac{d^2}{(d^1)^TGd^2}. \tag{9.99}$$

若(9.99)成立，$\det D=0$，显然$|\det D|$不是最大．现在设$0<|c|<1$．用下式定义新向量$\hat{d}^1$：

$$\hat{d}^1=\frac{d^1-cd^2}{[1-(c)^2]^{1/2}}, \tag{9.100}$$

定义新矩阵$\hat{D}$：

$$\hat{D}=[\hat{d}^1,\ d^2,\ \cdots,\ d^n]. \tag{9.101}$$

容易验证$\hat{d}^1$满足(9.95)，而且

$$|\det \hat{D}|=[1-(c)^2]^{-1/2}|\det D|>|\det D|. \tag{9.102}$$

也即$|\det D|$并未达到它的最大值．

反之，设d^1，…，d^n与$\bar{d}^1$，…，$\bar{d}^n$是不同的两组向量，它们都关于G共轭而且已按(9.95)规范化．于是存在一个$n\times n$实矩阵$P=[p_{ij}]$，使得

$$\bar{d}^i=\sum_{j=1}^n p_{ij}d^j,\quad i=1,\ \cdots,\ n. \tag{9.103}$$

设$\bar{D}$是以向量$\bar{d}^i$为列作出的矩阵，那么

$$\det \bar{D}=\det D\det P. \tag{9.104}$$

因此，我们必须证明$|\det P|=1$．

现在

$$(\bar{d}^i)^TG\bar{d}^k=\sum_{j=1}^n\sum_{l=1}^n p_{ij}p_{kl}(d^j)^TGd^l \tag{9.105}$$

$$=\sum_{j=1}^n\sum_{l=1}^n p_{ij}p_{kl}\delta_{jl}=\sum_{j=1}^n p_{ij}p_{kj}=\delta_{ik}, \tag{9.106}$$

这里，当$r=s$时$\delta_{rs}=1$，而当$r\neq s$时$\delta_{rs}=0$．从(9.106)得知，矩阵P的行构成一组n个规范的正交向量，因此P是一个正交阵，由此推出$|\det P|=1$． ▌

让我们应用上述结果于修改的Powell法，提出这个做法的目的在于避免搜索方向线性相关的可能性．假定有n个线性独立的搜索方向Δ_1^k，…，Δ_n^k，并假定

$$\Delta_j^k = \mu_j^k d_j^k, \quad j=1, \cdots, n, \tag{9.107}$$

因而 d_j^k 满足(9.95). 因此

$$(\Delta_j^k)^T G \Delta_j^k = (\mu_j^k)^2. \tag{9.108}$$

回想

$$\Delta_{n+1}^k = (t_n^k - t_0^k) = \sum_{j=1}^{n} \theta_j^* \Delta_j^k = \mu_{n+1}^k d_{n+1}^k, \tag{9.109}$$

这里 μ_{n+1}^k 与 d_{n+1}^k 选得使 d_{n+1}^k 满足规范化条件(9.95). 设矩阵 D^k 用下式给定:

$$D^k = [d_1^k, \cdots, d_n^k], \tag{9.110}$$

并以 D^{k+1} 记用 d_{n+1}^k 替换 D^k 的第 m 列 d_m^k 之后所得的矩阵. 现在由(9.107)及(9.109),

$$d_{n+1}^k = \sum_{j=1}^{n} \frac{\theta_j^* \mu_j^k}{\mu_{n+1}^k} d_j^k. \tag{9.111}$$

因此

$$D^{k+1} = \left[d_1^k, \cdots, \frac{\theta_1^* \mu_1^k d_1^k}{\mu_{n+1}^k} + \cdots + \frac{\theta_m^* \mu_m^k d_m^k}{\mu_{n+1}^k} + \cdots \right.$$
$$\left. + \frac{\theta_n^* \mu_n^k d_n^k}{\mu_{n+1}^k}, \cdots, d_n^k \right]. \tag{9.112}$$

根据线性代数知道，当把行列式中一列的某个倍数加到另一列时行列式值不变，当用常数 c 乘矩阵中一列的诸元素时，相应的行列式也被乘以 c, 因此从(9.112)得到

$$\det D^{k+1} = \frac{\theta_m^* \mu_m^k}{\mu_{n+1}^k} \det D^k. \tag{9.113}$$

为了逐渐有 n 个相互共轭的方向，定理 9.3 提示我们应选取指标 m, 使得 $\det D^{k+1}/\det D^k$ 的绝对值有最大的增加; 也即

$$|\theta_m^* \mu_m^k| = \max_{j=1,\cdots,n} |\theta_j^* \mu_j^k|. \tag{9.114}$$

注意到若

$$\left|\frac{\theta_m^* \mu_m^k}{\mu_{n+1}^k}\right| < 1, \tag{9.115}$$

则当第 m 个搜索方向 Δ_m^k 被代之以 Δ_{n+1}^k 时, $|\det D^{k+1}/\det D^k|$ 之值不会增加. 对于由(9.31)给出的二次函数 f, 可以指出有

$$f(t_{j-1}^k)-f(t_j^k)=b^T(-\theta_j^*\Delta_j^k)+\frac{1}{2}(2t_j^k-\theta_j^*\Delta_j^k)^TQ(-\theta_j^*\Delta_j^k). \tag{9.116}$$

因为 t_j^k 选得使 $f(t_{j-1}^k+\theta_j\Delta_j^k)$ 在该处最小，所以应有

$$(b+Qt_j^k)^T\Delta_j^k=0. \tag{9.117}$$

由(9.107)及(9.108)得

$$f(t_{j-1}^k)-f(t_j^k)=\frac{1}{2}(\theta_j^*\mu_j^k)^2. \tag{9.118}$$

因此(9.114)中表达式取极大的指标 m，对应于在这个算法的第 k 阶段中产生最大的函数值减少的搜索方向，正象(9.84)所给出的。容易证明

$$f(t_0^k)-f(t_0^{k+1})=\frac{1}{2}(\alpha_{n+1}^*\mu_{n+1}^k)^2, \tag{9.119}$$

从而不改变搜索方向的条件，正如(9.115)所给定的，可以改写为

$$|\alpha_{n+1}^*|<\left[\frac{f(t_0^k)-f(t_0^{k+1})}{f(t_{m-1}^k)-f(t_m^k)}\right]^{1/2}, \tag{9.120}$$

这正是(9.86)．显然，如果从一组线性独立的方向出发，随后的搜索方向将保持为独立的，而且按 Powell 行列式度量，它们将趋向于关于矩阵 Q 是共轭的．

如果要极小化的函数不是二次的，就没有固定的矩阵 G，因而就不能规范化方向 d_j^k，也就不能保证通过条件(9.120)可使 $|\det D^k|$ 的值不减小．虽然上述修改的 Powell 方法在极小化一般函数时是十分成功的[14]，但是为了保证搜索方向的线性独立性，条件将稍加修改．为了规范化，我们可以选取任何固定的矩阵 G，最简单的选取显然是 $G=I$（单位阵）．于是，对于 $k=1, 2, \cdots$，我们要求

$$(d_j^k)^Td_j^k=1, \quad j=1, \cdots, n, \tag{9.121}$$

而

$$\mu_j^k=\|\Delta_j^k\|, \quad j=1, \cdots, n. \tag{9.122}$$

方程(9.113)变成

$$\det D^{k+1}=\frac{\theta_m^*\|\Delta_m^k\|}{\|\Delta_{n+1}^k\|}\det D^k. \tag{9.123}$$

现在假定

$$|\det D^k|\geqslant\varepsilon, \tag{9.124}$$

这里 ε 是一个正常数．那么被 Δ_{n+1}^k 替换的方向的指标 m 将使得

$$|\theta_m^*|\,\|\Delta_m^k\|=\|t_m^k-t_{m-1}^k\|=\max_{j=1,\cdots,n}\|t_j^k-t_{j-1}^k\|, \tag{9.125}$$

而且如果

$$|\det D^{k+1}|=\frac{\|t_m^k-t_{m-1}^k\|}{\|\Delta_{n+1}^k\|}|\det D^k|<\varepsilon, \tag{9.126}$$

那么第 $k+1$ 阶段的搜索方向将保持与第 k 阶段相同．这个算法属于 Zangwill[17]，可概述如下：

给定 n 个线性独立的搜索方向 Δ_1^1, $\cdots$, Δ_n^1、一个出发点 $x_B^0\in R^n$ 和一个正数 ε，使得

$$|\det D^1|\geqslant\varepsilon. \tag{9.127}$$

这方法的第 k 阶段由下列步骤构成：

1．与 Powell 原来的方法一样找 t_j^k, $j=1$, $\cdots$, $n+1$，及 x_B^k.

2．找指标 m 使 $\|t_j^k-t_{j-1}^k\|$ $(j=1, \cdots, n)$ 达到最大.

3．若

$$|\det D^{k+1}|=\frac{\|t_m^k-t_{m-1}^k\|}{\|\Delta_{n+1}^k\|}|\det D^k|\geqslant\varepsilon, \tag{9.128}$$

取

$$\Delta_j^{k+1}=\Delta_j^k,\quad j=1,\ \cdots,\ n,\ \ j\neq m, \tag{9.129}$$

$$\Delta_m^{k+1}=\Delta_{n+1}^k. \tag{9.130}$$

若(9.128)不成立，取

$$\Delta_j^{k+1}=\Delta_j^k,\quad j=1,\ \cdots,\ n, \tag{9.131}$$

这时有

$$\det D^{k+1}=\det D^k. \tag{9.132}$$

再进入第 $k+1$ 阶段．算法的结束条件可以与 Powell 方法一样.

我们现在证明，对于连续可微严格伪凸函数，上述方法收敛到它的极小值点．应注意这些函数的全体比具有正定矩阵 Q 的二次

函数类要大得多. 证明基于 Zangwill[17]（他考虑严格凸函数）和 Daniol[4] 的工作.

首先，我们有下面直接的结果.

引理 9.4

对所有 k，搜索方向 $d_1^k, \cdots, d_n^k$ 是线性独立的.

【证明】 对 $k=1$，由假设可知引理是真的. 对 $k=2, 3, \cdots$，由步骤 3 我们有 $|\det D^k| \geqslant \varepsilon$. ▌

在下面讨论中，设 K（有时带有上标，例如 K^1）记自然数的一个子序列. 记号 $K^1 \subset K$ 表示 K^1 是 K 的一个子序列. 假定我们写 $\{p^k\}$, $k \in K$，它表示通过取 $k \in K$ 来构成 p^k 的一个子序列. 于是 $\{p^{k+1}\}$, $k \in K$，表明对每个 $k \in K$ 加 1 而构成的 p^k 的子序列.

引理 9.5

对任何子序列 K，有一个 $K^1 \subset K$，使得

$$\lim_{k \in K^1} \{d_j^k\} = d_j^0, \quad j=1, \cdots, n. \tag{9.133}$$

方向 d_j^0 是线性独立的，而且 $(d_j^0)^T d_j^0 = 1$, $j=1, \cdots, n$.

【证明】 因为 d_j^k 用 (9.121) 规范化，它们被限制在一个紧致集上. 因此有一个 $K^1 \subset K$，使得 (9.133) 成立而且 $(d_j^0)^T d_j^0 = 1$. 另外设

$$\lim_{k \in K^1} \{|\det D^k|\} = |\det[d_1^0, \cdots, d_n^0]| = |\det D^0|. \tag{9.134}$$

因为对所有的 k 有 $|\det D^k| \geqslant \varepsilon$，所以 $|\det D^0| \geqslant \varepsilon$，因而 d_j^0 是线性独立的. ▌

现在回忆一下第 6 章关于强拟凸函数与严格伪凸函数的定义. R^n 上的一个实值函数 f 称为是强拟凸的，如果对任意两点 $x^1 \neq x^2$ 和任意的 $q_1>0$, $q_2>0$, $q_1+q_2=1$ 有

$$f(q_1x^1+q_2x^2) < \max[f(x^1), f(x^2)]. \tag{9.135}$$

它称为是严格伪凸的，如果它是可微的，而且对任意两点 $x^1 \neq x^2$,

$$(x^1-x^2)^T \nabla f(x^2) \geqslant 0 \text{ 蕴涵 } f(x^1) > f(x^2). \tag{9.136}$$

在第 6 章中也曾指出一个强拟凸函数在 R^n 上（或它的某个凸子集上）达到它的极小值的点不多于一个. 特别地，这个结论对于严格

伪凸函数也是成立的,因为它们也是强拟凸的.

定理 9.6

设 f 是 R^n 上一个连续的强拟凸函数,从任意点 x_B^0 出发,对它用上述方法进行极小化. 假定水平集

$$S(f,\ x_B^0)=\{x:x\in R^n,\ f(x)\leqslant f(x_B^0)\} \tag{9.137}$$

是有界的. 那么,当 $k\to\infty$ 时 $\{t_j^k\}$ 的任何一个极限点 t^0 也是 $\{t_l^k\}$ $(l\neq j)$ 当 $k\to\infty$ 时的一个极限点. 对于每一个这样的极限点,存在着 n 个线性独立的方向 $d_1^0,\ \cdots,\ d_n^0$,使得对一切 θ_j 有

$$f(t^0)\leqslant f(t^0+\theta_j d_j^0),\quad j=1,\ \cdots,\ n. \tag{9.138}$$

【证明】 对于任何子序列 K,可以找到一个 $K^1\subset K$,使得

$$\lim_{k\in K^1}\{t_j^k\}=t_j^0,\quad j=0,\ \cdots,\ n. \tag{9.139}$$

我们现在来说明 $t_j^0=t_{j+1}^0$. 从算法可知,对所有 $k=1,\ 2,\ \cdots$,

$$f(t_{j+1}^k)\leqslant f(t_j^k),\quad j=0,\ \cdots,\ n-1, \tag{9.140}$$

而且 $f(t_0^{k+1})\leqslant f(t_n^k)$.

因为序列 $\{f(t_j^k)\}$ 是单调的,

$$\lim_{k\in K^1}\{f(t_{j+1}^k)\}=\lim_{k\in K^1}\{f(t_j^k)\},\quad j=0,\ \cdots,\ n-1. \tag{9.141}$$

由 f 的连续性,得

$$\lim_{k\in K^1}\{f(t_{j+1}^k)\}=f(t_{j+1}^0), \tag{9.142}$$

$$\lim_{k\in K^1}\{f(t_j^k)\}=f(t_j^0). \tag{9.143}$$

对所有 $\overline{\theta}_{j+1}$,我们有

$$f(t_{j+1}^k)\leqslant f(t_j^k+\overline{\theta}_{j+1}d_{j+1}^k), \tag{9.144}$$

再由连续性,对所有的 $\overline{\theta}_{j+1}$ 有

$$f(t_{j+1}^0)=f(t_j^0)\leqslant f(t_j^0+\overline{\theta}_{j+1}d_{j+1}^0),\quad j=0,\ \cdots,\ n-1. \tag{9.145}$$

如果 $t_{j+1}^0\neq t_j^0$,那么函数 f 沿方向 d_{j+1}^0 在两个不同点达到极小. 因为 f 是强拟凸的,由(9.135)推出

$$f(q_1t_j^0+q_2t_{j+1}^0)<f(t_j^0)=f(t_{j+1}^0), \tag{9.146}$$

得到了矛盾,因此 $t_{j+1}^0=t_j^0$. 事实上,

$$t^0=t_1^0=t_2^0=\cdots=t_n^0. \tag{9.147}$$

因此，对所有 θ_j，

$$f(t^0)\leqslant f(t^0+\theta_j d_j^0),\quad j=1,\ \cdots,\ n \tag{9.148}$$

而由引理 9.5 可知 d_j^0 是线性独立的。 ▌

推论 9.7

假定定理 9.6 的假设成立，并设 f 还是连续可微的，则对 $\{t_j^k\}$ 的任何极限点 t^0 有 $\nabla f(t^0)=0$.

【证明】 由定理 9.6，对一切 θ_j 有

$$f(t^0)\leqslant f(t^0+\theta_j d_j^0),\quad j=1,\ \cdots,\ n, \tag{9.149}$$

或

$$f(t^0)\leqslant f(t^0)+\theta_j(d_j^0)^T\nabla f(t^0+\lambda\theta_j d_j^0), \tag{9.150}$$

这里 $\lambda\in[0,\ 1]$. 使用类似于证明定理 2.3 时用过的推理，可以断定

$$(d_j^0)^T\nabla f(t^0)=0,\quad j=1,\ \cdots,\ n. \tag{9.151}$$

因为 d_j^0 是线性独立的，因此必定有 $\nabla f(t^0)=0$. ▌

可以证明[4]，若在上述推论中加上假设：集 $\{x:\nabla f(x)=0\}$ 不包含闭联集（闭联集是一个闭集，它多于一点*，却又不能表示成两个非空不相交闭集之和集），则 $\{t_j^k\}$ 收敛. 因此如果由 $\nabla f(t^0)=0$ 可推出 t^0 是一个连续可微的强拟凸函数的唯一的极小值点（该条件显然为严格伪凸函数所满足），那么便完成了收敛性的证明. Zangwill[17] 已经详尽地阐述了这种类型的方法，并且发展了一个具有二次终结性质和上述收敛性质的类似算法. 因为 Zangwill 的算法类似于前面所述的算法，还因为从计算的观点看，它不象是优良的[14]，所以关于进一步细节请读者参考[71].

9.7 进一步的共轭方向型算法

另外两个不用导数的无约束极小化算法在此给出. 首先我们叙述并证明 Powell 的一个结果，它涉及到使搜索方向趋于相互共轭的一种新方法. 这个结果可以作为一个不用导数的新的极小化算法的基础. 第二个论题涉及到多维极小化的序列同时性分批型

* 译注：这一句是译者添加的.

搜索方法．这个方法可以看作是 Powell 的第一个方法对于具有平行算术处理器的计算机的一个推广．

下列结果属于 Powell[12]．

定理 9.8

设 $d^1, \cdots, d^n$ 是一组满足规范化条件 (9.95) 的向量，并设 $P=[p_{ij}]$ 是一个正交阵．用下式定义一组新向量：

$$\bar{d}^i=\sum_{j=1}^{n} p_{ij} d^j, \quad i=1, \cdots, n, \tag{9.152}$$

并应用规范化公式得

$$\hat{d}^i=\frac{\bar{d}^i}{[(\bar{d}^i)^T G \bar{d}^i]^{1/2}}, \tag{9.153}$$

其中 $G=[g_{jk}]$ 是一个对称正定阵．若以 D 及 $\hat{D}$ 分别表示以向量 d^i 及 $\hat{d}^i$ 作为列的矩阵，则

$$|\det \hat{D}| \geqslant |\det D|. \tag{9.154}$$

【证明】 从(9.152)我们有 $\bar{D}=DP^T$，其中 $\bar{D}$ 是以 $\bar{d}^i$ 为列构成的矩阵．于是

$$|\det \bar{D}|=|\det D|\,|\det P|=|\det D|, \tag{9.155}$$

$$|\det \hat{D}|=\frac{|\det \bar{D}|}{\Pi^{1/2}}, \tag{9.156}$$

这里

$$\Pi=[(\bar{d}^1)^T G \bar{d}^1][(\bar{d}^2)^T G \bar{d}^2]\cdots[(\bar{d}^n)^T G \bar{d}^n]. \tag{9.157}$$

我们必须证明 $\Pi \leqslant 1$．因为 G 是正定的，Π 的每一项都是正的．因此利用关于算术平均值与几何平均值的不等式，得到

$$\Pi \leqslant \left[\sum_{i=1}^{n} \frac{(\bar{d}^i)^T G \bar{d}^i}{n}\right]^n. \tag{9.158}$$

由(9.95)、(9.152)及 P 的正交性，我们得到

$$\sum_{i=1}^{n}(\bar{d}^i)^T G \bar{d}^i=\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} g_{jk}\bar{d}_j^i\bar{d}_k^i \tag{9.159}$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\sum_{m=1}^{n} g_{jk}p_{il}d_j^l p_{im}d_k^m \tag{9.160}$$

$$=\sum_{l=1}^{n}(d^l)^T G d^l=n. \tag{9.161}$$

▮

上述定理表明，如果我们用一个正交阵 P，把一组当前的搜索方向 $(d^1, \cdots, d^n)$ 通过 (9.152) 及 (9.153) 改变到一组新的方向 $(\hat{d}^1, \cdots, \hat{d}^n)$，那么这组新方向将至少同样"接近"共轭性，这里的"接近"用当前方向的行列式值来度量. 因此只要有一个适当的矩阵 G，就可以用(9.152)和(9.153)作为更新搜索方向的一个规则. 相当有趣的是，对于 $n=2$，Powell 找到了一个正交阵

$$P_{(2)}=\begin{bmatrix} \cos\frac{\pi}{4} & \sin\frac{\pi}{4} \\ -\sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{bmatrix}=\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, \qquad (9.162)$$

它把任意两个方向 d^1, d^2（已用对称正定阵 G 按(9.95)规范化）变换为关于 G 共轭的方向 $\hat{d}^1$, $\hat{d}^2$. 注意，$P_{(2)}$ 是独立于有关的方向及矩阵的.

例 9.7.1

设

$$G=\begin{bmatrix} 4 & 2 \\ 2 & 16 \end{bmatrix}, \quad d^1=\begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix}, \quad d^2=\begin{pmatrix} 0 \\ \frac{1}{4} \end{pmatrix}. \qquad (9.163)$$

向量 d^1 与 d^2 关于 G 是规范化的，但它们关于 G 不是共轭的. 现在设 $P_{(2)}$ 由(9.162)给出. 那么由(9.152)，

$$\bar{d}^1=\frac{1}{\sqrt{2}}d^1+\frac{1}{\sqrt{2}}d^2=\begin{pmatrix} \frac{1}{2\sqrt{2}} \\ \frac{1}{4\sqrt{2}} \end{pmatrix}, \qquad (9.164)$$

$$\bar{d}^2=\frac{-1}{\sqrt{2}}d^1+\frac{1}{\sqrt{2}}d^2=\begin{pmatrix} -\frac{1}{2\sqrt{2}} \\ \frac{1}{4\sqrt{2}} \end{pmatrix}, \qquad (9.165)$$

$$(\bar{d}^1)^T G\bar{d}^1=\frac{5}{4}, \quad (\bar{d}^2)^T G\bar{d}^2=\frac{3}{4}. \qquad (9.166)$$

因而，由(9.153)得

$$\hat{d}^1=\begin{pmatrix}\dfrac{1}{\sqrt{10}}\\[2ex]\dfrac{1}{2\sqrt{10}}\end{pmatrix},\quad \hat{d}^2=\begin{pmatrix}-\dfrac{1}{\sqrt{6}}\\[2ex]\dfrac{1}{2\sqrt{6}}\end{pmatrix}, \tag{9.167}$$

它们关于 G 是规范化的又是共轭的. ▌

Powell[12] 并未通过精确地陈述一个新算法来详细说明上述定理中的概念. 然而我们可以设想如何根据定理 9.8 来建立一个算法.

这样一个算法的成功看来将依赖于估计二阶导数和计算(9.153)中的分母. 对于由(9.31)给定的二次函数，显然有 $G=Q$. 对于一个一般的函数，G 必须是 Hesse 阵的一个估计. 假定我们用 f 在一点 $x^*\in R^n$ 周围的 Taylor 展开式作为任意实函数 f 的二次函数逼近：

$$f(x)\cong f(x^*)+(x-x^*)^T\nabla f(x^*)+\frac{1}{2}(x-x^*)^T\nabla^2 f(x^*)(x-x^*). \tag{9.168}$$

对于一个给定的点 x 及方向 d，设 x^* 是 $f(x+\theta d)$ 极小值点的猜测，也即 $x^*=x+\theta^* d$. 那末

$$d^T\nabla^2 f(x^*)d\cong\frac{2[f(x)-f(x^*)]}{(\theta^*)^2}. \tag{9.169}$$

这个关系式在某些迭代方式中可以用于(9.153). 这里可以假定，若 $\theta^*=0$，则 $d^T\nabla^2 f(x^*)d=1$.

至今尚没有基于上述定理的数值经验可供引用，不过我们相信这个结果是重要的，基于它的新方法最终将会出现.

正象在一维最优化的情形一样，我们用一个分批型搜索方法来结束本章，这个方法用同时进行一维最优化来推广 Powell[11] 及 Zangwill[17] 的算法. 这个由 Chazan、Miranker[3] 提出的算法是为使用于具有平行算术处理器的计算机[6]而建立的. 事实上，它生成一组共轭方向，并具有二次终结性质. 作者们已对严格凸函数证明了算法的收敛性.

我们对极小化 R^n 上一个无约束函数 f 的情况来叙述这个算法，然后再通过一个例子来说明．设 $u^1, \cdots, u^n$ 是 R^n 中 n 个单位长度的线性独立方向(通常就取坐标方向)，再设 $z_1^1, \cdots, z_{n-1}^1$ 是 R^n 中 $n-1$ 个方向，并假定已给定一个初始点 $t_0^1\in R^n$．希望选取 z_j^1 使之也是线性独立的．在第 k 轮迭代中，我们有方向 $z_1^k, \cdots, z_{n-1}^k$，一个点 t_0^k 及一个整数 r, $1\leqslant r\leqslant n$．设

$$z_n^k=u^r. \tag{9.170}$$

定义 n 个点如下：

$$p_j^k=t_0^k+\sum_{i=1}^{j}z_i^k,\quad j=1,\cdots,n. \tag{9.171}$$

然后从点 $p_1^k, \cdots, p_n^k$ 出发沿着相同的方向 z_1^k 同时进行 n 个的极小化．这就是说，找 $\theta_1^*, \cdots, \theta_n^*$ 使得

$$f(p_j^k+\theta_j^* z_1^k)=\min_{\theta_j} f(p_j^k+\theta_j z_1^k),\quad j=1,\cdots,n. \tag{9.172}$$

下一步是取

$$t_0^{k+1}=p_1^k+\theta_1^* z_1^k, \tag{9.173}$$

$$z_j^{k+1}=z_{j+1}^k+(\theta_{j+1}^*-\theta_j^*)z_1^k,\quad j=1,\cdots,n-1. \tag{9.174}$$

若 $r<n$，就把指标 r 增加 1；否则，取 $r=1$．进入第 $k+1$ 轮迭代．这个方法的少数步骤在图 9.6 中说明．

这个算法的主要目标是用同时线性搜索来减少计算时间，因为 Powell 方法及有关算法的最费时间的部分是一维极小化步骤．为了极小化一个 R^n 中的二次函数所需的线性搜索数是 (n^2) 阶的．如果每个一维极小化需一个单位时间，那么由于算法的序列性，对于二次终结就要有 $(n)^2$ 个单位时间．另一方面，上述 Chazan-Miranker 的同时-序列算法为了完成同一任务只需 n 个单位时间．

例 9.7.2

为了对 9.5.1 的二次函数说明 Chazan-Miranker 算法，寻找下列函数的极小值点：

$$f(x)=\frac{3}{2}(x_1)^2+\frac{1}{2}(x_2)^2-x_1x_2-2x_1. \tag{9.175}$$

假定我们从 $t_0^1=(-2, 4)$ 出发．设

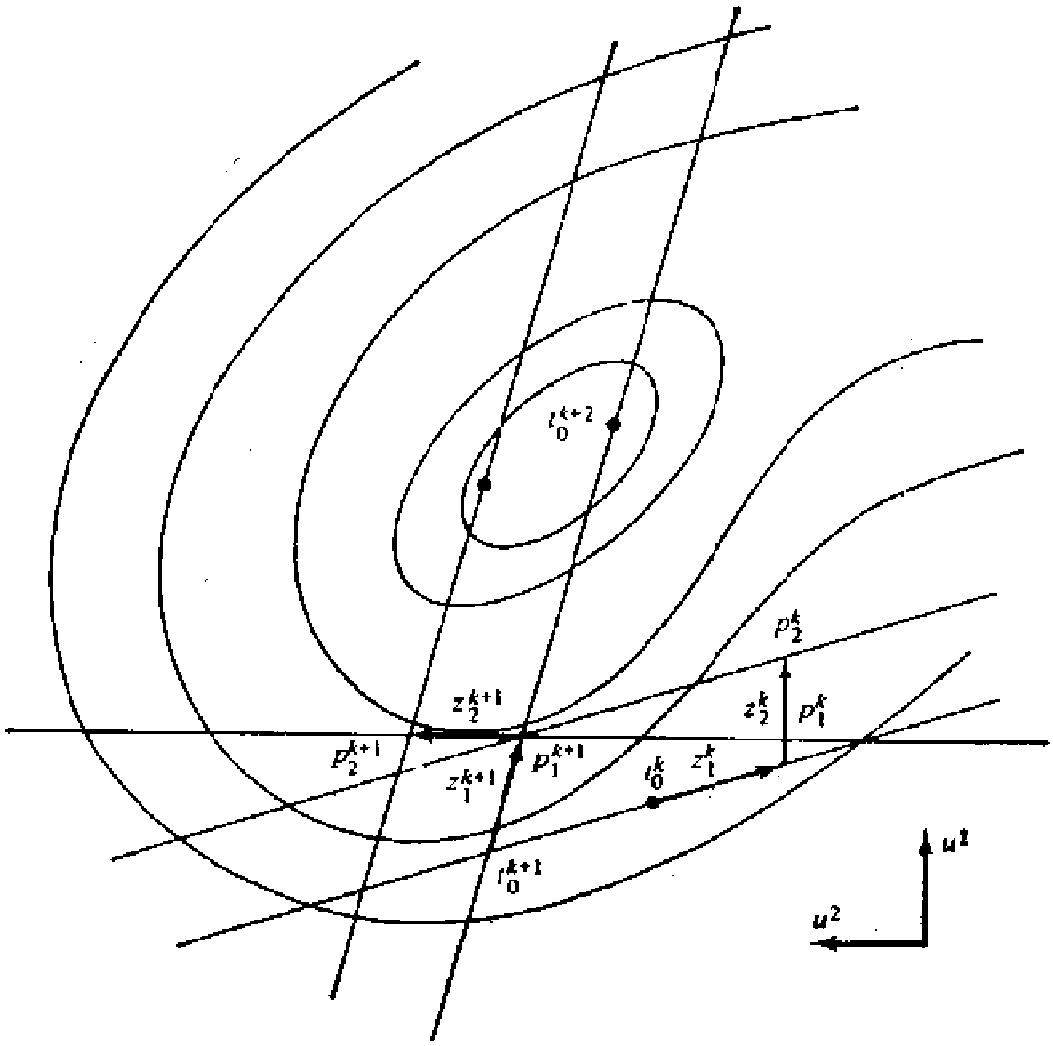


图 9.6 同时-序列的非梯度算法

$$u^1=\begin{pmatrix}1\\0\end{pmatrix},\quad u^2=\begin{pmatrix}0\\1\end{pmatrix},\quad z_1^1=\begin{pmatrix}-1\\2\end{pmatrix}, \tag{9.176}$$

而 $r=1$，那么 $z_2^1=u^1$. 由(9.171)，

$$p_1^1=\begin{pmatrix}-2\\4\end{pmatrix}+\begin{pmatrix}-1\\2\end{pmatrix}=\begin{pmatrix}-3\\6\end{pmatrix}, \tag{9.177}$$

$$p_2^1=\begin{pmatrix}-2\\4\end{pmatrix}+\begin{pmatrix}-1\\2\end{pmatrix}+\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}-2\\6\end{pmatrix}. \tag{9.178}$$

现在从 p_1^1 及 p_2^1 沿 z_1^1 方向同时进行极小化，这两个极小化为

$$\min_{\theta_1} f(p_1^1+\theta_1 z_1^1)=\frac{3}{2}(-3-\theta_1)^2+\frac{1}{2}(6+2\theta_1)^2 \\ -(-3-\theta_1)(6+2\theta_1)-2(-3-\theta_1), \tag{9.179}$$

$$\min_{\theta_2} f(p_2^1+\theta_2 z_1^1)=\frac{3}{2}(-2-\theta_2)^2+\frac{1}{2}(6+2\theta_2)^2-(-2-\theta_2)(6+2\theta_2)-2(-2-\theta_2), \tag{9.180}$$

其最优解分别为 $\theta_1^*=-\frac{35}{11}$, $\theta_2^*=-\frac{30}{11}$, 因此

$$t_0^1=\begin{pmatrix}-3\\6\end{pmatrix}-\frac{35}{11}\begin{pmatrix}-1\\2\end{pmatrix}=\begin{pmatrix}\frac{2}{11}\\-\frac{4}{11}\end{pmatrix}, \tag{9.181}$$

$$z_2^1=\begin{pmatrix}1\\0\end{pmatrix}+\left(-\frac{30}{11}+\frac{35}{11}\right)\begin{pmatrix}-1\\2\end{pmatrix}=\begin{pmatrix}\frac{6}{11}\\\frac{10}{11}\end{pmatrix}. \tag{9.182}$$

令 $r=2$, 我们得到 $z_2^2=u^2$. 第一轮迭代到此结束. 接着进入第二轮迭代, 得到

$$p_1^2=\begin{pmatrix}\frac{2}{11}\\-\frac{4}{11}\end{pmatrix}+\begin{pmatrix}\frac{6}{11}\\\frac{10}{11}\end{pmatrix}=\begin{pmatrix}\frac{8}{11}\\\frac{6}{11}\end{pmatrix}, \tag{9.183}$$

$$p_2^2=\begin{pmatrix}\frac{2}{11}\\-\frac{4}{11}\end{pmatrix}+\begin{pmatrix}\frac{6}{11}\\\frac{10}{11}\end{pmatrix}+\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}\frac{8}{11}\\\frac{17}{11}\end{pmatrix}. \tag{9.184}$$

两个同时极小化是

$$\min_{\theta_1} f(p_1^2+\theta_1 z_1^2)=\frac{3}{2}\left(\frac{8}{11}+\frac{6}{11}\theta_1\right)^2+\frac{1}{2}\left(\frac{6}{11}+\frac{10}{11}\theta_1\right)^2-\left(\frac{8}{11}+\frac{6}{11}\theta_1\right)\left(\frac{6}{11}+\frac{10}{11}\theta_1\right)-2\left(\frac{8}{11}+\frac{6}{11}\theta_1\right), \tag{9.185}$$

$$\min_{\theta_2} f(p_2^2+\theta_2 z_1^2)=\frac{3}{2}\left(\frac{8}{11}+\frac{6}{11}\theta_2\right)^2+\frac{1}{2}\left(\frac{17}{11}+\frac{10}{11}\theta_2\right)^2-\left(\frac{8}{11}+\frac{6}{11}\theta_2\right)\left(\frac{17}{11}+\frac{10}{11}\theta_2\right)-2\left(\frac{8}{11}+\frac{6}{11}\theta_2\right), \tag{9.186}$$

分别在$\theta_1^*=\frac{1}{2}$和$\theta_2^*=-\frac{17}{22}$达到极小值，因此

$$t_0^3=\begin{pmatrix}\frac{8}{11}\\[2mm]\frac{6}{11}\end{pmatrix}+\left(\frac{1}{2}\right)\begin{pmatrix}\frac{6}{11}\\[2mm]\frac{10}{11}\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}, \qquad (9.187)$$

我们看到，经过$n=2$轮迭代，f的整体最优点确实已达到．读者可以验证，方向z_1^1和z_1^2关于(9.175)的矩阵Q是共轭的． ∎

这个算法的二次终结性质的证明留给读者；对于更一般函数的收敛性的证明可以在[3]中找到．

练　习

9.A. 最早的n变量无约束最优化方法或许是所谓的交替方向法(或称轴向松弛法)．在这种方法中，一个实函数$f(x)$ $(x\in R^n)$的极小值点通过从任意点出发依次沿着坐标轴方向搜索而找到．画出这个方法的框图，它采用了上一章的某个一维搜索方法．对下列函数试验这个算法：

(i) $$f_1(x)=(x_1-3)^2+2(x_2-2)^2, \qquad (9.188)$$

从$x^0=(0, 0)$出发．

(ii) $$f_2(x)=(x_1)^2+106(x_2)^2+10x_1x_2, \qquad (9.189)$$

从$x^0=(4, 3)$出发．

求数α和β，使得变换

$$w_1=x_1+\alpha x_2, \quad w_2=\beta x_2 \qquad (9.190)$$

对于f_2得出

$$f_2(w)=(w_1)^2+(w_2)^2. \qquad (9.191)$$

讨论这种变换对于交替方向法的效率所起的作用．

9.B. 使用单纯形方法极小化下列函数：

(i) $$f_1(x)=(x_1-3)^2+(x_2-2)^2+(x_1+x_2-4)^2, \qquad (9.192)$$

从下列点开始：

$$x^0=(0, 8),\quad x^1=(0, 9),\quad x^2=(1, 9). \qquad (9.193)$$

用$\alpha=1$，$\beta=\frac{1}{2}$，$\gamma=2$．画出这个二次函数的等值线，并且画出这个算法的进程．

(ii) $$f_3(x)=(-6-x_1-x_2)^2+(2-3x_1-3x_2-x_1x_2)^2, \tag{9.194}$$
从下列点开始：
$$x^0=(-4, 6),\quad x^1=(-4, 7),\quad x^2=(-3, 6). \tag{9.195}$$
这个函数（它在 $x^*=(0, 0)$ 有一个极小值点）是电子工程最优化问题中出现的目标函数的一种形式.

对这两个函数应用交替方向法，并比较两种搜索方法对于这些试验函数的效率.

9.C. 画出模式搜索法的框图，并在一个计算机上把它编成程序. 对前面练习中的函数试验你的程序. 假如你还没有接触到一个计算机，就将模式搜索法应用于极小化下列函数的问题：
$$f(x)=(x_1)^2+(x_2)^2-3x_1-x_1x_2+3. \tag{9.196}$$
从 $x_B^0=(0, 0)$ 出发搜索，开始取 $d_1=1$, $d_2=1$. 完成勘探与模式阶段. 画出这个函数的等值线，并说明这个算法的行程. 把你的结果与例 9.3.1 比较.

9.D. 假定在旋转方向法的第 k 个勘探阶段的末了，有着 $p(>0)$ 个方向，其 α_j 为 0，说明为什么在这个场合找新方向的 Gram-Schmidt 正交化方法将会失效. 若 $p(>0)$ 个 α_j 为 0，得到一组新方向的一种可能的方法是：重排 u^j 使得前 $(n-p)$ 个具有非零的 α_j 并如 (9.17) 那样定义. 对于 $j=1, \cdots, n-p$ 进行 Rosenbrock 所建议的正交化并得到新的 ζ^j, $j=1, \cdots, n-p$. 定义其余的 ζ^j 与刚完成的勘探阶段中的相同. 试说明这样得到的 n 个方向向量是相互正交的.

9.E. 把旋转方向法应用于极小化 R^2 上的一个函数，假定其第 k 阶段的最后步骤的结果如下表所列：

经 过 的 点	函 数 值
(1, 8)	25.0
(2, 9)	26.0
(0, 9)	22.5
(−0.5, 8.5)	21.0
(−3.5, 11.5)	24.8
(−2, 7)	23.6

什么是 $\zeta^{1,k}$ 及 $\zeta^{2,k}$? 什么将是搜索方向 $\zeta^{1,k+1}$, $\zeta^{2,k+1}$?

9.F. 使用旋转方向法于函数 (9.192) 及 (9.196)，分别从 $x_B^0=(0, 8)$ 及 (0, 0) 出发，首先选 $\zeta^1=(1, 0)$、$\zeta^2=(0, 1)$，取 $\alpha=3$、$\beta=0.5$，画出这个

算法的行程并与单纯形法及模式搜索法比较.

9.G. 以 Davies-Swann-Campey (DSC) 方法代替 Rosenbrock 旋转方向法重做练习 9.F.

9.H. 证明公式(9.45)及(9.46)产生线性独立的相互共轭的方向. 并证明, 若 $s^1, \cdots, s^k$ 是非零向量, 且关于对称正定阵 A 相互共轭, 则它们是线性独立的.

9.I. 关于下列矩阵, 找出一组相互共轭的方向:

$$\text{(a)}\ G_1=\begin{bmatrix} 1 & 2 \\ 2 & 9 \end{bmatrix}, \qquad \text{(b)}\ G_2=\begin{bmatrix} 2 & -2 & 0 \\ -2 & 3 & -1 \\ 0 & -1 & 6 \end{bmatrix}. \tag{9.197}$$

假定 s^1, s^2 与 s^1, s^3 是关于一个 3×3 正定对称阵 G 的共轭方向, 是否 s^2, s^3 也是关于 G 共轭的?

9.J. 给定 R^n 上一个具有正定阵 Q 的二次函数 f, 假定 $\theta_1^*=\theta_2^*=\cdots=\theta_n^*=0$, 这里的最优步长 θ_j^* 通过下式定义:

$$f(x^0+\theta_j^*\Delta^j)=\min_{\theta_j} f(x^0+\theta_j\Delta^j), \quad j=1, \cdots, n, \tag{9.198}$$

其中 Δ^j 是正交方向. 证明 x^0 是 f 的极小值点. 对于什么类型的函数, 可以推广这个结果?

9.K. 证明, 例 9.5.1 中的方向 Δ_3^1 与 Δ_3^2 关于(9.68)的二次函数所对应的矩阵 Q 是共轭的.

9.L. 写出不用导数极小化实函数的计算机程序, 以实现 9.5 节的 Powel 法, 它用带有外推(不用划界方法)的二次逼近线性搜索, 在开始时采用平行于坐标轴的搜索方向. 当两个相继的迭代的终点之差小于某个容许范围时, 收敛准则就满足了. 为了检验这个程序, 对下列问题:

$$\min f_1(x)=10(x_1)^2+(x_2)^2-60x_1-8x_2+106, \tag{9.199}$$

从不同点出发用上述方法求解. 在所有场合, 同一个最优解经过 6 个线性搜索后由计算机打印了出来. 另一方面, 当问题

$$\min f_2(x)=3(x_1)^2+(x_2)^2-2x_1x_2-4x_1 \tag{9.200}$$

用同样方法求解时, 最优解要 9 个线性搜索后才打印出来. 试解释为了得到这两个问题各自的最优解所需线性搜索数目的这个差别. 假定计算机绝对精确, 估计为解这些问题所需计算函数值的总数.

9.M. 把例 9.6.1 继续进行两个阶段. 计算在每个阶段中用 (9.95) 规范化的搜索方向的行列式, 并且注意这些方向是否逐渐近于共轭.

9.N. 把 Zangwill 算法用来极小化例 9.6.1 的二次函数, 计算三个阶段.

9.O. 设 G 是习题 9.I 中的 3×3 矩阵，选一个 3×3 正交阵 P 及 R^3 中的三个方向. 用矩阵 G 与 P 得到一组新的方向，它们至少象原来的方向一样近于相互共轭(在定理 9.8 的意义下). 从一组三个关于 G 相互共轭的方向出发重复这个计算.

9.P. 画出 Chazan-Miranker 方法的框图. 应用这算法于二次函数(9.188)，从原点出发，并点出它的行程.

参 考 文 献

1. Box, M. J., "A Comparison of Several Current Optimization Methods, and the Use of Transformations in Constrained Problems," *Computer J.*, **9**, 67–77 (1966).
2. Box, M. J., D. Davies, and W. H. Swann, *Non-Linear Optimization Techniques*, I.C.I. Monograph No. 5, Oliver & Boyd, Edinburgh, 1969.
3. Chazan, D., and W. L. Miranker, "A Nongradient and Parallel Algorithm for Unconstrained Minimization," *SIAM J. Control*, **8**, 207–217 (1970).
4. Daniel, J. W., *The Approximate Minimization of Functionals*, Prentice-Hall, Englewood Cliffs, N.J., 1971.
5. Hooke, R., and T. A. Jeeves, "Direct Search Solution of Numerical and Statistical Problems," *J. Assoc. Comp. Mach.*, **8**, 212–221 (1961).
6. Miranker, W. L., "A Survey of Parallelism in Numerical Analysis," *SIAM Review*, **13**, 524–547 (1971).
7. Nelder, J. A., and R. Mead, "A Simplex Method for Function Minimization," *Computer J.*, **7**, 308–313 (1965).
8. Noble, B., *Applied Linear Algebra*, Prentice-Hall, Englewood Cliffs, N.J., 1969.
9. Palmer, J. R., "An Improved Procedure for Orthogonalising the Search Vectors in Rosenbrock's and Swann's Direct Search Optimization Methods," *Computer J.*, **12**, 69–71 (1969).
10. Pierre, D. A., *Optimization Theory with Applications*, John Wiley & Sons, New York, 1969.
11. Powell, M. J. D., "An Efficient Method for Finding the Minimum of a Function of Several Variables without Calculating Derivatives," *Computer J.*, **7**, 155–162 (1964).
12. Powell, M. J. D., "Unconstrained Minimization Algorithms without Computation of Derivatives," Report T. P. 483, A.E.R.A. Harwell, United Kingdom, April 1972.
13. Rosenbrock, H. H., "An Automatic Method for Finding the Greatest or Least Value of a Function," *Computer J.*, **3**, 175–184 (1960).

14. SARGENT, R. W. H., "Minimization without Constraints," in *Optimization and Design*, M. Avriel, M. J. Rijckaert, and D. J. Wilde (Eds.), Prentice-Hall, Englewood Cliffs, N.J., 1973.

15. SPENDLEY, W., G. R. HEXT, and F. R. HIMSWORTH, "Sequential Application of Simplex Designs in Optimisation and Evolutionary Operation," *Technometrics*, **4**, 441–461 (1962).

16. SWANN, W. H., "Report on the Development of a New Direct Search Method of Optimization," I.C.I. Central Instrument Lab. Research Note 64/3, 1964.

17. ZANGWILL, W. I., "Minimizing a Function without Calculating Derivatives", *Computer J.*, **10**, 293–296 (1967).

第 10 章

二阶导数法、最速下降法和共轭梯度法

在关于利用导数的无约束极小化方法的以下两章中，先从两个经典算法——Newton 法和 Cauchy 最速下降法开始讨论. Newton 法是 8.1 节所述算法对多维的拓广. 对于多变量的一般非线性函数的极小化，这两个方法都不够满意. 比如，Newton 法可以不收敛于所寻求的极小值点，而最速下降法的收敛速率很慢. 然而，为了理解从它们经过各种修改而得到的、在实践中认为非常成功的那些方法，它们是重要的. 在介绍了综合上面两个经典算法的 Goldstein-Price 方法之后，再转向共轭梯度法.

最初由解线性方程组而发展起来的共轭梯度法的根本原理是，大多数非线性函数在其极小值点的邻域中能用一个二次函数来相当好地近似. 这样，任何有效的极小化迭代方法对二次函数的极小化应是特别有效. 这里，我们回到共轭方向的概念，并用它导出共轭梯度法. 本章以这方法的收敛性的讨论来结束. 共轭梯度法的一个有趣而不那么好理解的特性是它揭示了更为复杂和精致的变尺度法的许多重要性质. 变尺度法被认为是最好的无约束极小化算法，将在下一章中讨论.

10.1 Newton 型法和最速下降法

假设实值函数 f 在 R^n 上连续可微. 从第 2 章知道，f 在某点 $x^*\in R^n$ 达到极小的一个必要(有时是充分的)条件是

$$\nabla f(x^*)=0. \tag{10.1}$$

如果我们有关于这个函数的足够信息，知道(10.1)的解是一个极小值点，那就可以去解上面那个具有 n 个变量由 n 个一般是非线性方程组成的方程组. 恰与第 8 章中讨论的 $n=1$ 的情形相同，这

里求(10.1)的解的经典方法也是解方程组的 Newton 法．为了应用 Newton 法，必须假定 f 至少二次连续可微，即除梯度向量 $\nabla f(x)$ 之外，还需要在每一点 $x\in R^n$ 处的 n 阶 Hesse 阵 $\nabla^2 f(x)$．将 ∇f 的每一个分量在点 x^k 附近展开(线性化)，并令这个线性函数为零：

$$\frac{\partial f(x^k)}{\partial x_j}+\sum_{i=1}^{n}\frac{\partial^2 f(x^k)}{\partial x_i\,\partial x_j}(x_i-x_i^k)=0,\quad j=1,\ \cdots,\ n,\tag{10.2}$$

或用向量记号表示为

$$\nabla f(x^k)+\nabla^2 f(x^k)(x-x^k)=0.\tag{10.3}$$

假定 $\nabla^2 f(x^k)$ 是非异的，则能解上述关于 x 的线性方程组．设 x^{k+1} 是解，我们得到

$$x^{k+1}=x^k-[\nabla^2 f(x^k)]^{-1}\nabla f(x^k).\tag{10.4}$$

Newton 法就是用(10.4)进行迭代．在每次迭代中，f 在 x^k 的梯度和 Hesse 逆阵必须算出．实际上，我们不直接去求 Hesse 逆阵而是去解线性方程组(10.2)．在适当的假定下，类似于对一维情形已介绍过的那样，Newton 法收敛于(10.1)的一个解．有关收敛性的透彻讨论，读者可参考 Goldstein[16]，Ortega、Rheinboldt[25] 或 Ostrowski[27]．

当有大批变量时，即使对一个处处存在正定 Hesse 阵的性质很好的函数，计算函数值和导数值，特别是求逆阵的过程都是耗时很多且规模浩大的工作．如果 Hesse 阵近于奇异，局面则更坏．将 Newton 法用于一般函数时，经常出现的另一个困难是，由(10.4)得到的 x^{k+1} 不能保证函数值在每次迭代中都减少．换言之，可能有 $f(x^{k+1})>f(x^k)$，且由(10.4)进行迭代而生成的点列可能收敛于 f 的鞍点或极大值点．

有另一种考察 Newton 法的途径．假设在 x^k 附近把 f 展开成 Taylor 级数，以作出它在某点 x 的二阶近似：

$$f(x)\cong f(x^k)+(x-x^k)^T\nabla f(x^k)+\frac{1}{2}(x-x^k)^T\nabla^2 f(x^k)(x-x^k).\tag{10.5}$$

倘若 $\nabla^2 f(x^k)$ 正定，则由 (10.4) 给出的点 x^{k+1} 是这个二阶近似的极小值点. 用 Newton 法时大多数困难发生在下述情形中: $\nabla^2 f$ 在 x^k 非正定，或者由 (10.4) 得到的点 x^{k+1} 不靠近 x^k 以致于 f 在 x^k 附近的二阶近似在 x^{k+1} 处无效. 部分的补救办法已提出. 比如,可用一个**限步 Newton 公式**

$$x^{k+1}=x^k-\theta[\nabla^2 f(x^k)]^{-1}\nabla f(x^k) \tag{10.6}$$

来代替(10.4),其中 θ 选择得使 $f(x^{k+1})<f(x^k)$，而且在这算法的某些形式中,θ 选择得使 $f(x)$ 沿这个 Newton 方向取极小, 这一步可用第 8 章中的线性搜索来进行.

关于对一般函数求极小的 Newton 法，一个有趣的改进已由 Goldfeld、Quandt、Trotter[14] 提出. 他们算法的基础是，在第 k 次迭代中, 对由(10.5)给出的 f 的二阶近似在以 x^k 为中心的闭球上求极小. 球的大小由这个算法来调整，并以 f 的二阶近似在球中保持相当好为条件，取得尽可能的大. 这个算法后来改进为以一椭球代替球形区域而在其上对二阶近似求极小. 这个椭球的形状和方位在每次迭代中修改.

基于少量小型的试验问题, 几位作者报告说, 他们的算法很有生命力，总的执行情况可与不用导数的 Powell 法(第 9 章)相比. 然而看来, 如同原先的 Newton 法一样, 这些算法的效率也随着变量数目的增加而迅速减小，其主要原因是在每次迭代中需要求 Hesse 逆阵, 或者等价地, 需要去解一个线性方程组.

有一种情形，Newton 法变得极为简单且在一次迭代中收敛. 这种情形出现在 f 是由下式给出的二次函数时:

$$f(x)=a+b^T x+\frac{1}{2}x^T Q x, \tag{10.7}$$

其中 Q 为正定阵. 注意，$\nabla^2 f(x)=Q$，即 Hesse 阵是一个常数矩阵. 设 x^0 是 R^n 中任意一点, x^* 是 f 的极小值点, 则有

$$\nabla f(x^0)=b+Qx^0, \tag{10.8}$$

$$0=b+Qx^*. \tag{10.9}$$

综合这两个方程, 就得到

$$x^* = x^0 - Q^{-1}\nabla f(x^0). \tag{10.10}$$

比较(10.10)与(10.4)可见，从任一点 x^0 出发，用 Newton 法只要进行一步就达到这二次函数的极小值点.

例 10.1.1

再考察例 9.5.2 中用过的二次函数 $f(x)=(x_1-x_2+x_3)^2+(-x_1+x_2+x_3)^2+(x_1+x_2-x_3)^2$. 用简单的代数处理可将它变形为

$$f(x)=\frac{1}{2}x^TQx, \quad x\in R^3, \tag{10.11}$$

其中，Q 及其逆阵 Q^{-1} 由下式给出:

$$Q=\begin{bmatrix} 6 & -2 & -2 \\ -2 & 6 & -2 \\ -2 & -2 & 6 \end{bmatrix}, \quad Q^{-1}=\begin{bmatrix} \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{4} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{4} \end{bmatrix}. \tag{10.12}$$

设 $x^0=\left(\frac{1}{2}, 1, \frac{1}{2}\right)$, 则

$$\nabla f(x^0)=Qx^0=\begin{bmatrix} 6 & -2 & -2 \\ -2 & 6 & -2 \\ -2 & -2 & 6 \end{bmatrix}\begin{pmatrix} \frac{1}{2} \\ 1 \\ \frac{1}{2} \end{pmatrix}=\begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix}, \tag{10.13}$$

且由(10.10)得到精确的极小值点

$$x^*=\begin{pmatrix} \frac{1}{2} \\ 1 \\ \frac{1}{2} \end{pmatrix}-\begin{bmatrix} \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{4} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{4} \end{bmatrix}\begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{10.14}$$

∎

Newton 法利用迭代格式去求这样一点，在这一点上求极小的函数的梯度向量为零. 与梯度向量有关的一个应用涉及到寻找函

数值减小的方向．在第 4 章中，我们曾定义可微函数 f 的方向导数为

$$Df(x^0;\ y)=y^T\nabla f(x^0)=\lim_{t\to 0}\frac{f(x^0+ty)-f(x^0)}{t}. \tag{10.15}$$

现在考察所有这样的向量(方向) $y\in R^n$，使得对给定的点 $x^0\in R^n$ 有

$$y^T\nabla f(x^0)<0. \tag{10.16}$$

于是由(10.15)可得，对充分小的正数 t 有

$$f(x^0+ty)<f(x^0). \tag{10.17}$$

换言之，如果我们要找 f 在 R^n 上的极小值点而在某一点 $x^0\in R^n$ 上 f 的梯度不为零，那么，沿着满足(10.16)的 y 方向作一个充分小的移动将导致函数值减小．方向导数 $Df(x^0;y)$，实际上可以度量 f 的值在 x^0 沿 y 方向的瞬时增加（如果 $Df(x^0;\ y)>0$）或瞬时减小(如果 $Df(x^0;\ y)<0$)．因此，可在具有某有界长度的所有方向 y 中，比如在 $\|y\|\leqslant 1$ 的所有方向中，找一个特殊方向，使得沿此方向，在符合 $\nabla f(x^0)\neq 0$ 的给定点 x^0，f 的值能最速下降．于是，有下述非线性规划问题：

$$\min_y y^T\nabla f\ (x^0)=\sum_{j=1}^n\frac{\partial f(x^0)}{\partial x_j}y_j, \tag{10.18}$$

受限制于

$$\|y\|=\left\{\sum_{j=1}^n(y_j)^2\right\}^{1/2}\leqslant 1. \tag{10.19}$$

要求读者证明，这问题的最优解是

$$y^*=-\frac{\nabla f(x^0)}{\|\nabla f\ (x^0)\|}. \tag{10.20}$$

这样，函数值最速下降的方向是负梯度方向．首先由 Cauchy[3] 导出的最速下降法可描述如下：给定一点 $x^0\in R^n$，对 $k=0,\ 1,\ \cdots$，计算点列

$$x^{k+1}=x^k-\theta_k^*\nabla f(x^k), \tag{10.21}$$

其中 $\theta_k^*>0$，且满足

$$f(x^k-\theta_k^*\nabla f(x^k))=\min_{\theta_k\geqslant 0}f(x^k-\theta_k\nabla f(x^k)). \tag{10.22}$$

在 Cauchy 最速下降法中，要求得 f 沿负梯度方向的整体极小值点. Curry[6] 对此方法作了修改，他选择 θ_k^* 是 f 沿方向 $-\nabla f(x^k)$ 的第一个逗留点，即 θ_k 取最小的正值时得到的那个逗留点.

由(10.20)给出的最速下降的方向向量，在改变(10.19)的模约束时不是不变的. 例如，假定代替 Euclid 模条件而要求

$$y^T A y \leqslant 1, \tag{10.23}$$

其中 A 为一正定对称阵，我们极小化 $y^T\nabla f(x^0)$ 使它受限制于(10.23). 仍请读者证明，这种情形的最优解 y 由下式给出：

$$y^A = \frac{-A^{-1}\nabla f(x^0)}{[(\nabla f(x^0))^T A^{-1}\nabla f(x^0)]^{1/2}}. \tag{10.24}$$

如果在每次迭代中令 $A=\nabla^2 f(x^k)$，倘若 Hesse 阵为正定，我们得到在利用非 Euclid 模的最速下降法与本节开始时叙述的 Newton 法之间有趣的相似. 然而，我们在下面仅涉及利用 Euclid 模的最速下降法. Greenstadt[17] 研究了梯度方法应用于二次函数极小化时，不同的矩阵 A 对于相对效率(将在下面定义)的影响. 由这一研究，也可试着去得出有关非二次函数的结论.

假设根据(10.24)选一个由 x^0 出发的搜索方向 y^A，沿此方向极小化一个二次函数. 其最优值点记为 x^A，则有

$$\begin{aligned} f(x^0)-f(x^A) &= \frac{[(y^A)^T\nabla f(x^0)]^2}{2(y^A)^T\nabla^2 f(x^0)y^A} \\ &= \frac{[(\nabla f(x^0))^T A^{-1}\nabla f(x^0)]^2}{2(\nabla f(x^0))^T A^{-1}\nabla^2 f(x^0)A^{-1}\nabla f(x^0)}. \end{aligned} \tag{10.25}$$

令 $A=\nabla^2 f(x^0)$，这搜索方向即变为 Newton 法的方向，记为 y^N. 在 Newton 方向上进行极小化，得到一点 x^N，并且

$$f(x^0)-f(x^N)=\frac{1}{2}(\nabla f(x^0))^T[\nabla^2 f(x^0)]^{-1}\nabla f(x^0). \tag{10.26}$$

定义相对效率 η 为

$$\eta=\frac{f(x^0)-f(x^A)}{f(x^0)-f(x^N)}, \tag{10.27}$$

定义矩阵 M 为

$$M = A^{-1/2}\nabla^2 f(x^0)A^{-1/2}, \qquad (10.28)$$

其中 $A^{-1/2}$ 是 A^{-1} 的对称正定“平方根”——即 $A^{-1/2}A^{-1/2}=A^{-1}$，可得

$$\frac{4\mu}{(1+\mu)^2}\leqslant\eta\leqslant 1, \qquad (10.29)$$

其中 μ 是 M 的最大特征值与最小特征值之比，也称它为 M 的条件数. η 以 1 为上界表达了一个前已建立的事实：从任一点 x^0 出发，用 Newton 法只要单独一步，便可达到具有正定 Hesse 阵的二次函数的极小值点.

当应用于有狭长“深谷”的函数时，不等式(10.29)也能约略地启示出最速下降法($A=I$)的不好的执行情况. 对这样的函数，M 的条件数将是大的，从而最速下降法作出的进展很慢. 这种现象称为花边现象(见图 10.1).

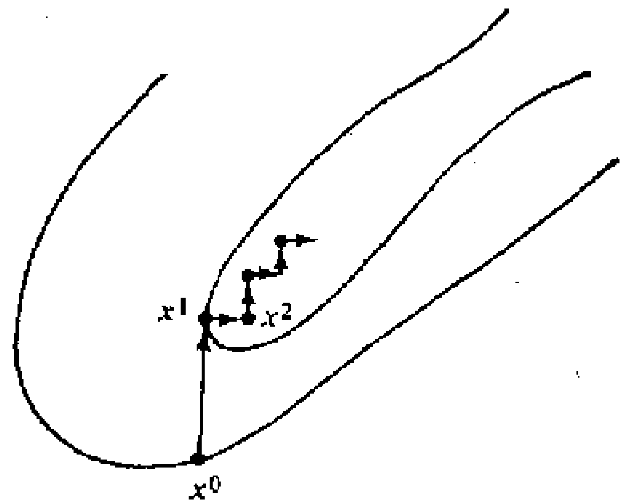


图 10.1 由最速下降法产生的花边现象

让我们看看关于最速下降法的收敛性有些什么可说的. 请读者回想一下，$\hat{x}\in R^n$ 称为序列 $\{x^k\}\subset R^n$ 的一个凝聚点，如果对每一个正数 ε，存在一个自然数 $K(\varepsilon)$，使得对 $k=K(\varepsilon)$ 有 $x^k\in N_\varepsilon(\hat{x})$*，其中

$$N_\varepsilon(\hat{x})=\{x:\|x-\hat{x}\|<\varepsilon\}. \qquad (10.30)$$

现在可以叙述并证明下面的定理.

* 译注：原文此处为对一切 $k\geqslant K(\varepsilon)$ 有 $x^k\in N_\varepsilon(\hat{x})$. 这与极限点的概念混淆.

定理 10.1[15]

设 f 是 R^n 上的一个实值函数. 给定 $x^0 \in R^n$, $f(x^0)=\alpha$. 假定水平集

$$S(f,\ \alpha)=\{x: x\in R^n,\ f(x)\leqslant \alpha\} \tag{10.31}$$

有界, 且 f 在 $S(f,\ \alpha)$ 的凸包上连续可微. 令 $\{x^k\}$ 是由最速下降法产生的点列, 则 $\{x^k\}$ 的每一个凝聚点 $\hat{x}$ 满足 $\nabla f(\hat{x})=0$.

【证明】 因为 $S(f,\ \alpha)$ 有界, 由 Bolzano-Weierstrass 定理[2] 可知 $\{x^k\}$ 至少存在一个凝聚点. 令 $\hat{x}$ 是这样一个凝聚点, 并设 $\nabla f(\hat{x})\neq 0$. 于是, 存在一个 $\hat{\theta}^*>0$, 使 $f(\hat{x}-\hat{\theta}\nabla f(\hat{x}))$ 达到极小; 即存在一个 $\delta>0$, 使得 $f(\hat{x})=f(\hat{x}-\hat{\theta}^*\nabla f(\hat{x}))+\delta$. 因此 $\hat{x}-\hat{\theta}^*\nabla f(\hat{x})$ 是 $S(f,\ \alpha)$ 的一个内点.

对任意 x^k, 由中值定理可写出

$$\begin{aligned} f(x^k-\hat{\theta}^*\nabla f(x^k)) &= f(\hat{x}-\hat{\theta}^*\nabla f(\hat{x})) \\ &\quad +(\nabla f(\xi^k))^T[x^k-\hat{x}+\hat{\theta}^*(\nabla f(\hat{x})-\nabla f(x^k))], \end{aligned} \tag{10.32}$$

其中 $\xi^k=\hat{x}-\hat{\theta}^*\nabla f(\hat{x})+\lambda[(x^k-\hat{x})-\hat{\theta}^*(\nabla f(x^k)-\nabla f(\hat{x}))]$, λ 是 $(0,\ 1)$ 中的某个值. 由 Bolzano-Weierstrass 定理也能推得, 存在 $\{x^k\}$ 的一个子序列 $\{x^{k_m}\}$ 收敛于 $\hat{x}$. 故可知 $\{\nabla f(\xi^{k_m})\}$ 收敛于 $\nabla f(\hat{x}-\hat{\theta}^*\nabla f(\hat{x}))$, 且 $\{(x^{k_m}-\hat{x})-\hat{\theta}^*(\nabla f(x^{k_m})-\nabla f(\hat{x}))\}$ 收敛于零. 对充分大的 k_m, 向量 ξ^{k_m} 属于 $S(f,\ \alpha)$ 的凸包, 且

$$f(x^{k_m}-\hat{\theta}^*\nabla f(x^{k_m}))\leqslant f(\hat{x}-\hat{\theta}^*\nabla f(\hat{x}))+\frac{\delta}{2}=f(\hat{x})-\frac{\delta}{2}. \tag{10.33}$$

令 $\theta^*_{k_m}$ 为 $f(x^{k_m}-\theta_{k_m}\nabla f(x^{k_m}))$ 的极小值点. 因为 $\{f(x^{k_m})\}$ 单调下降且收敛于 $f(\hat{x})$, 得

$$\begin{aligned} f(\hat{x}) &< f(x^{k_m}-\theta^*_{k_m}\nabla f(x^{k_m}))\leqslant f(x^{k_m}-\hat{\theta}^*\nabla f(x^{k_m})) \\ &\leqslant f(\hat{x})-\frac{\delta}{2}, \end{aligned} \tag{10.34}$$

与假设矛盾. 因此, 必须 $\nabla f(\hat{x})=0$. 】

这个定理能够推广, 只需用下列公式代替在最速下降法中的 (10.21):

$$x^{k+1}=x^k-\bar{\theta}_k B(x^k)\nabla f(x^k), \qquad (10.35)$$

其中 $B(x)$ 是一个 $n\times n$ 的正定阵，其元素为 x 的连续函数，$\bar{\theta}_k\geqslant 0$ 是一个数，它由上面提到过的 Curry 规则确定．这样，一些利用非 Euclid 模的"最速下降"方法在定理 10.1 的假设条件下也收敛．前已指出，有关最速下降法的计算经验通常使人失望，除了函数具有近乎球形等值面的情形之外，都收敛得慢．由 Akaike[1]，Forsythe[13]，Kantorovich-Akilov[19] 所作的理论研究表明，最速下降法的收敛速率对二次函数是线性的，并且一般说来，对非二次函数不可能指望有更高的速率．

更特殊地，可不失一般性地假定二次函数 f 为 $f(x)=\frac{1}{2}x^TQx$，其中 Q 是一个 $n\times n$ 正定对称阵，其最小和最大的特征值分别为 λ_1 和 λ_n．这个函数的整体极小值点当然在原点．Forsythe[13] 证明了，从任意点 $x^0\in R^n$ 开始，对于以最速下降法生成的点，存在依赖于 x^0 的常数 c_1 和 c_2，满足

$$0<c_1\leqslant\frac{f(x^{k+1})}{f(x^k)}\leqslant c_2\leqslant\left(\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}\right)^2<1, \quad k=0,\ 1,\ \cdots. \qquad (10.36)$$

考察一下相继的搜索方向就可以对这样慢的收敛速率作出启发性的解释．假设在第 k 次迭代得到点 x^k，$\nabla f(x^k)\neq 0$．我们求函数 $f(x^k-\theta_k\nabla f(x^k))$ 关于 θ_k 的极小并令 $x^{k+1}=x^k-\theta_k^*\nabla f(x^k)$．由简单的微分可证明 $(\nabla f(x^k))^T\nabla f(x^{k+1})=0$；即相继的搜索方向互相垂直．在 R^2 中，这意味着我们仅仅沿着平行于开始的两个方向进行搜索．在高于二维的情形，这算法实际上也相似，其搜索方向渐近地趋于两个固定向量[13]．

我们叙述 Goldstein 和 Price[16] 的一个极小化算法来结束这一节．这算法可看作是把 Newton 法和最速下降法结合起来并消除了这两者的某些不希望有的特性而得到的一个方法．在下面的讨论中，δ 和 r 是正数，$\delta<\frac{1}{2}$，x^0 是 R^n 中任意一点，I_i 是 $n\times n$ 单位阵 I 的第 i 列．f 是至少定义在 (10.31) 给出的水平集上的实

值函数．求 f 极小的 Goldstein-Price 算法的第 k 次迭代由下面的计算组成．计算矩阵 $Q(x^k)$，它的第 i 列由下式给出：

$$Q_i(x^k)=\frac{\nabla f(x^k+\beta_k I_i)-\nabla f(x^k)}{\beta_k},\quad i=1,\ \cdots,\ n,\tag{10.37}$$

其中，$\beta_0=r$，而

$$\beta_k=r\|\phi(x^{k-1})\|,\quad k=1,\ 2,\ \cdots.\tag{10.38}$$

这里，$\phi(x^{k-1})$ 按以下两种途径给出：

$$\phi(x^k)=\nabla f(x^k)\tag{10.39}$$

用于 $k=0$ 时，或者 $Q(x^k)$ 是奇异时，或者下式成立时，

$$(\nabla f(x^k))^T[Q(x^k)]^{-1}\nabla f(x^k)\leqslant 0.\tag{10.40}$$

否则，令

$$\phi(x^k)=[Q(x^k)]^{-1}\nabla f(x^k).\tag{10.41}$$

定义

$$g(x^k,\ \theta)=\frac{f(x^k)-f(x^k-\theta\phi(x^k))}{\theta(\nabla f(x^k))^T\phi(x^k)}.\tag{10.42}$$

如果 $g(x^k,\ 1)<\delta$，令 θ_k 满足

$$\delta\leqslant g(x^k,\ \theta_k)\leqslant 1-\delta;\tag{10.43}$$

否则令 $\theta_k=1$．下一个点 x^{k+1} 由下式给出：

$$x^{k+1}=x^k-\theta_k\phi(x^k).\tag{10.44}$$

如果 $\|\nabla f(x^k)\|<\varepsilon$，则迭代终止．这里 ε 是预定的小的正数．

在用数字例子来说明这个算法之前，作几点评注是适宜的．矩阵 $Q(x^k)$ 是 f 在 x^k 的 Hesse 阵 $\nabla^2 f(x^k)$ 的一个近似．这一近似的得到仅用到梯度，而不需计算二阶导数．在这算法中的搜索方向，或者是最速下降方向 $-\nabla f(x^k)$，或者是某一个近似的 Newton 方向 $-[Q(x^k)]^{-1}\nabla f(x^k)$．最速下降方向用于以下几种情形：首次迭代、$Q(x^k)$ 是奇异的因而不可能求逆（实际上，一旦 $Q(x^k)$ 接近奇异时，我们就选择最速下降方向），或者近似的 Newton 方向不能引起瞬时的函数减小，即

$$-(\nabla f(x^k))^T[Q(x^k)]^{-1}\nabla f(x^k)\geqslant 0.\tag{10.45}$$

如果取 $\theta_k=1$（"完全 Newton 步"）而得到 $g(x^k,\ 1)<\delta$，例如当

$f(x^k-\phi(x^k))>f(x^k)$时就是这种情形，则选择一个较小的步长使得$g(x^k, \theta_k)$为正且有异于0的下界．考察一个具有正定Hesse阵的二次函数f，就能很好地理解(10.43)中的上界和下界．在此情形下，$Q(x^k)$成为常数矩阵，即f的精确Hesse阵．在负梯度方向上第一次移动之后，下一个方向将根据Newton法确定，即$\phi(x^1)$由(10.41)给出．在x^1近旁将$g(x^1, \theta)$展开成Taylor级数，经过一些简单的代数处理就得到

$$g(x^1, \theta)=1-\frac{\theta}{2}, \tag{10.46}$$

而$\theta=1$的完全Newton步也满足(10.43)．

例 10.1.2

考察用Goldstein-Price算法极小化二次函数

$$f(x)=\frac{3}{2}(x_1)^2+\frac{1}{2}(x_2)^2-x_1x_2-2x_1, \tag{10.47}$$

这个函数在例9.5.1中曾用Powell方法求过极小．令$r=\frac{1}{2}$，$\delta=\frac{1}{8}$，并假设从$x^0=(-2, 4)$出发．f的梯度由下式给出：

$$\nabla f(x)=\begin{pmatrix} 3x_1-x_2-2 \\ x_2-x_1 \end{pmatrix}, \tag{10.48}$$

而

$$\nabla f(x^0)=\begin{pmatrix} -12 \\ 6 \end{pmatrix}. \tag{10.49}$$

矩阵$Q(x^0)$的列由下式给出：

$$Q_1(x^0)=\begin{pmatrix} 3 \\ -1 \end{pmatrix}, \quad Q_2(x^0)=\begin{pmatrix} -1 \\ 1 \end{pmatrix}. \tag{10.50}$$

注意，因f是二次的，矩阵Q恰是f的常数Hesse阵．第一个搜索方向由$\phi(x^0)=\nabla f(x^0)$给出．现在计算$g(x^0, 1)$：

$$g(x^0, 1)=\frac{f(x^0)-f(x^0-\nabla f(x^0))}{(\nabla f(x^0))^T\nabla f(x^0)} \tag{10.51}$$

$$=\frac{26-152}{180}=-\frac{126}{180}<\frac{1}{8}. \tag{10.52}$$

找一个 θ_0 使其满足(10.43). 试取 $\theta_0=\frac{1}{10}$, 得到

$$g\left(x^0,\ \frac{1}{10}\right)=\frac{f(x^0)-f\left(x^0-\left(\frac{1}{10}\right)\nabla f(x^0)\right)}{\left(\frac{1}{10}\right)(\nabla f(x^0))^T\nabla f(x^0)} \tag{10.53}$$

$$=\frac{26-11.06}{18}=\frac{14.94}{18}, \tag{10.54}$$

便成立

$$\frac{1}{8}<\frac{14.94}{18}<\frac{7}{8}. \tag{10.55}$$

所以 $x^1=(-0.8,\ 3.4)$. 注意到 $Q(x^1)=Q(x^0)=Q$, 算得

$$Q^{-1}=\begin{bmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{3}{2}\end{bmatrix}. \tag{10.56}$$

下一个搜索方向将是 Newton 方向

$$\phi(x^1)=Q^{-1}\nabla f(x^1)=\begin{bmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{3}{2}\end{bmatrix}\begin{pmatrix}-7.8\\ 4.2\end{pmatrix}=\begin{pmatrix}-1.8\\ 2.4\end{pmatrix}, \tag{10.57}$$

$$g(x^1,\ 1)=\frac{11.06-(-1)}{(-7.8)(-1.8)+(4.2)(2.4)}=\frac{1}{2}>\frac{1}{8}. \tag{10.58}$$

因此 $x^2=(1,\ 1)$且 $\nabla f(x^2)=0$. 即整体极小值点已经达到. 这一算法的行程以图 10.2 来说明. ∎

在这个例子中, 算法经过有限次迭代后结束. Goldstein 和 Price[16] 对广泛的一类 R^n 上严格凸函数建立了以超线性速率趋于整体极小的收敛性. 他们所做的有限的数值经验表明, 这一方法较之本章后面和下一章中将叙述的其他一些极小化算法要有利. 然而, 这是单就计算函数值和梯度值的次数所作的比较, 没有考虑在每次迭代中为了求 $Q(x^k)$的逆阵所需要的额外计算.

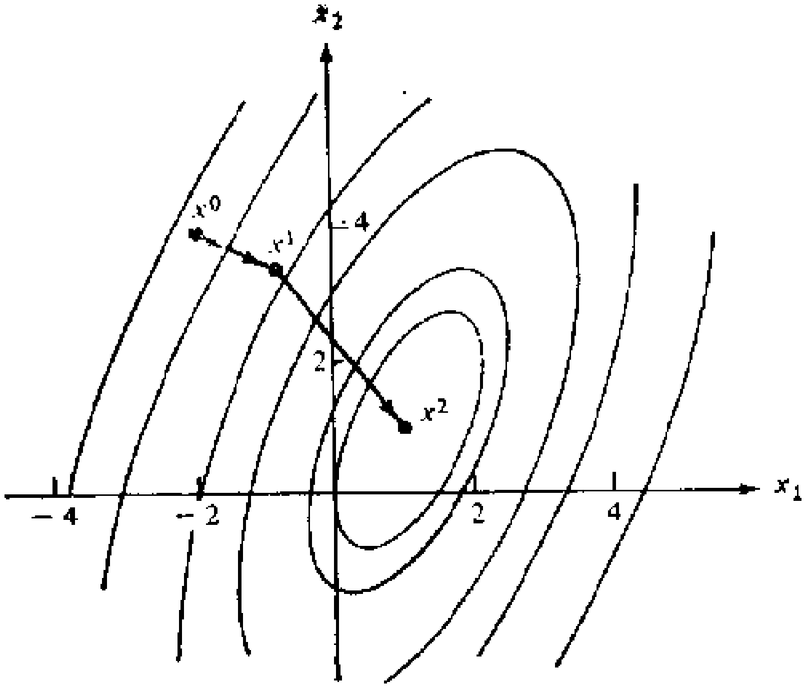


图 10.2 Goldstein-Price 算法

可以证明，矩阵 $Q(x^k)$ 按某种意义趋于求极小的函数的真正 Hesse 阵 $\nabla^2 f(x^k)$. 因此，可以试着去导出一些相似的算法. 在这些算法中，有一矩阵序列在极限过程中趋于 Hesse 逆阵 $[\nabla^2 f(x^k)]^{-1}$，而不必在每次迭代中求逆阵. 这种方法将在下一章中详尽地讨论.

另一些利用二阶导数的 Newton 型极小化方法在 Murray 的 [23] 中有叙述. Ritter[31] 提出了一个 Newton 型的算法，对于这个算法，在与 Goldstein-Price 方法相同的假设条件下能够证明超线性收敛性. Ritter 算法的收敛性（不一定是超线性的）在定理 10.1 的更为一般的条件下也已建立.

10.2 共轭梯度法

求 R^n 上一个实函数 f 的极小值点的迭代方法可以描述为：从初始点 x^0 移动到新点 x^1，再到 x^2，等等，作一系列移动，后继的点由下述关系式给出：

$$x^k = x^{k-1} + \alpha_k z^k. \tag{10.59}$$

这里 x^{k-1} 是当前点，z^k 是移动的方向向量，而 α_k 是步长. 假设方

向 z^k 已给定，并选择 α_k 使函数 f 沿 z^k 达到极小．令

$$F(\alpha_k)=f(x^{k-1}+\alpha_k z^k), \tag{10.60}$$

于是，在 F 的极小值点 α_k^* 处有

$$\frac{dF(\alpha_k^*)}{d\alpha_k}=(z^k)^T\nabla f(x^{k-1}+\alpha_k^* z^k)=(z^k)^T\nabla f(x^k)=0. \tag{10.61}$$

假定 f 是前面给出的一个二次函数：

$$f(x)=a+b^Tx+\frac{1}{2}x^TQx, \tag{10.62}$$

其中 Q 为一 $n\times n$ 正定对称阵．在这种情形下，f 在任意两点的梯度有如下关系：

$$\nabla f(x^k)=\nabla f(x^{k-1})+Q(x^k-x^{k-1}). \tag{10.63}$$

如果 $x^k=x^{k-1}+\alpha_k^* z^k$，则由(10.59)、(10.61)和(10.63)得到 α_k^* 的显式表示

$$\alpha_k^*=-\frac{(z^k)^T\nabla f(x^{k-1})}{(z^k)^TQz^k}. \tag{10.64}$$

在两点的函数值之间有下列关系式：

$$f(x^k)=f(x^{k-1})+(x^k-x^{k-1})^T\nabla f(x^{k-1})+\frac{1}{2}(x^k-x^{k-1})^TQ(x^k-x^{k-1}), \tag{10.65}$$

由(10.59)至(10.65)，得到

$$f(x^{k-1})-f(x^k)=\frac{[(z^k)^T\nabla f(x^{k-1})]^2}{2(z^k)^TQz^k}. \tag{10.66}$$

因已假定 Q 为正定，故这等式的右边对 $z^k\neq 0$ 是非负的，而当 z^k 不垂直于 $\nabla f(x^{k-1})$时是正的．在后一种情形下，因为 $f(x^{k-1})>f(x^k)$，故称算法为**下降方法**．我们只要求$(z^k)^T\nabla f(x^{k-1})\neq 0$，所以由(10.64)推得，若$(z^k)^T\nabla f(x^{k-1})>0$，则 $\alpha_k^*<0$(在最速下降法的讨论中，我们知道前面的数量积为负，它蕴涵 $\alpha_k^*>0$)．

除了下降的极小化方法之外，也希望有一种迅速收敛的算法，或更好些，希望有一种算法，把它用于求二次函数极小时只需有限步就能终结．因为一般非线性函数在极小值点的邻域中能用一个二次函数相当好地逼近，所以在一般函数的情形下，为了迅速收

敛，似乎也希望有二次终结性质．由上一章的结果，特别是定理9.1可知，若对 $k=1,\cdots,n$，搜索方向 z^k 关于 Q 相互共轭，则所得的点 x^n 是二次函数 f 的精确极小值点．共轭方向可按如下方式选择．

假设由一点 $x^0\in R^n$ 开始，选择

$$z^1=-\nabla f(x^0). \tag{10.67}$$

下一点是

$$x^1=x^0+\alpha_1^*z^1, \tag{10.68}$$

其中 α_1^* 由(10.64)给定．求出 $\nabla f(x^1)$ 的值并取

$$z^2=-\nabla f(x^1)+\beta_{11}z^1, \tag{10.69}$$

其中 β_{11} 是选定的一个数，使 z^1 与 z^2 关于 Q 共轭．因此

$$(z^1)^TQz^2=(z^1)^TQ(-\nabla f(x^1)+\beta_{11}z^1)=0, \tag{10.70}$$

从而

$$\beta_{11}=\frac{(z^1)^TQ\nabla f(x^1)}{(z^1)^TQz^1}. \tag{10.71}$$

现在从 x^1 沿着方向 z^2 移动到新的点 x^2，并计算 $\nabla f(x^2)$．新的方向 z^3(倘若 $n\geqslant 3$)应共轭于 z^1 和 z^2，要做到这一点，只要按下式取 z^3：

$$z^3=-\nabla f(x^2)+\beta_{21}z^1+\beta_{22}z^2, \tag{10.72}$$

其中 β_{21} 和 β_{22} 选择得使 $(z^1)^TQz^3=(z^2)^TQz^3=0$．一般地，有

$$z^{k+1}=-\nabla f(x^k)+\sum_{j=1}^{k}\beta_{kj}z^j,\quad k=0,\cdots,n-1. \tag{10.73}$$

使用这一公式的困难在于系数 β_{kj} 是 Q 的函数，若企图将(10.73)用于非二次函数，就必须计算 Hesse 阵，这是一个讨厌的运算．因此，我们将说明如何生成这些方向而不明显地利用 Q．下面的结果在以后是有用的．

定理 10.2

设 f 由(10.62)给出，并设 m 个非零向量 $z^1,\cdots,z^m$，$z^j\in R^n$，$m\leqslant n$ 关于 Q 相互共轭．由一点 $x^0\in R^n$ 出发，沿 $z^1,z^2,\cdots,z^m$ 分别移动到 $x^1,x^2,\cdots,x^m$，使得

$$(z^j)^T\nabla f(x^j)=0,\quad j=1,\ \cdots,\ m, \tag{10.74}$$

则

$$(z^j)^T\nabla f(x^m)=0,\quad j=1,\ \cdots,\ m. \tag{10.75}$$

【证明】 对 $j=m$，结论是平凡的. 从(10.63)有

$$\nabla f(x^m)=\nabla f(x^j)+Q(x^m-x^j),\quad j=0,\ \cdots,\ m-1, \tag{10.76}$$

且

$$x^j=x^{j-1}+\alpha_j^* z^j,\quad j=1,\ \cdots,\ m. \tag{10.77}$$

其中 α_j^* 的选择使得(10.74)成立. 再利用(10.77)，递推地可得

$$x^m-x^j=\sum_{i=j+1}^{m}\alpha_i^* z^i,\quad j=0,\ \cdots,\ m-1. \tag{10.78}$$

代入(10.76)后有

$$\nabla f(x^m)=\nabla f(x^j)+\sum_{i=j+1}^{m}\alpha_i^* Qz^i,\quad j=0,\ \cdots,\ m-1, \tag{10.79}$$

$$(z^j)^T\nabla f(x^m)=(z^j)^T\nabla f(x^j)+\sum_{i=j+1}^{m}\alpha_i^*(z^j)^TQz^i,$$
$$j=1,\ \cdots,\ m-1. \tag{10.80}$$

由(10.74)，右边第一项为零，由共轭性，第二项也为零. 因此(10.75)成立. ▌

推论 10.3

在定理 10.2 中，若 $m=n$，则

$$\nabla f(x^n)=0, \tag{10.81}$$

即 x^n 是 f 的无约束极小值点.

【证明】 因 z^j 是线性无关的，由(10.75)，f 在 x^n 的梯度必须为零. ▌

用 γ^i 来记 f 在相继两点 x^i 和 x^{i-1} 的梯度之差，即

$$\gamma^i=\nabla f(x^i)-\nabla f(x^{i-1}),\quad i=1,\ 2,\ \cdots. \tag{10.82}$$

对二次函数 f，由(10.63)可得

$$\gamma^i=Q(x^i-x^{i-1}). \tag{10.83}$$

点 x^i 和方向 z^i 的选择使得(10.59)和(10.83)成立. 这样，

$$(\gamma^i)^Tz^j=\alpha_i^*(z^i)^TQz^j,\quad i=1,\ \cdots,\ n,\ j=1,\ \cdots,\ n. \tag{10.84}$$

假定选择 $z^1,\ \cdots,\ z^k,\ k\leqslant n$，使之关于 Q 相互共轭，就得到

$$(\gamma^i)^T z^j = 0, \quad i=1, \cdots, k, \quad j=1, \cdots, k, \quad i \neq j. \tag{10.85}$$

我们利用这个结果去求系数 β_{11} 的另一种表示式.

$$(\gamma^1)^T z^2 = (\nabla f(x^1) - \nabla f(x^0))^T [-\nabla f(x^1) + \beta_{11} \nabla f(x^0)] = 0, \tag{10.86}$$

从而

$$\beta_{11} = \frac{(\nabla f(x^1) - \nabla f(x^0))^T \nabla f(x^1)}{(\nabla f(x^1) - \nabla f(x^0))^T (-\nabla f(x^0))} = \frac{(\gamma^1)^T \nabla f(x^1)}{(\gamma^1)^T (-\nabla f(x^0))}. \tag{10.87}$$

既然由(10.61)和(10.67)知 $(z^1)^T \nabla f(x^1) = -(\nabla f(x^0))^T \nabla f(x^1) = 0$，且假定 $\nabla f(x^0) \neq 0$，所以得到 β_{11} 的另一种表示式

$$\beta_{11} = \frac{(\nabla f(x^1))^T \nabla f(x^1)}{(\nabla f(x^0))^T \nabla f(x^0)}. \tag{10.88}$$

点 x^2 由沿着共轭方向 z^1 和 z^2 进行极小化达到. 因此, 由定理 10.2 有

$$(z^1)^T \nabla f(x^2) = (z^2)^T \nabla f(x^2) = 0. \tag{10.89}$$

将(10.67)和(10.69)代入(10.89)，我们断定

$$(\nabla f(x^0))^T \nabla f(x^2) = 0, \tag{10.90}$$

$$(\nabla f(x^1))^T \nabla f(x^2) = 0. \tag{10.91}$$

现在我们可找到 (10.72) 中系数 β_{21} 和 β_{22} 的一个公式. 将方程 $(\gamma^1)^T z^3 = (\gamma^2)^T z^3 = 0$ 与(10.89)至(10.91)综合起来，就得到

$$\beta_{21} = 0, \quad \beta_{22} = \frac{(\nabla f(x^2))^T \nabla f(x^2)}{(\nabla f(x^1))^T \nabla f(x^1)}. \tag{10.92}$$

注意, β_{22} (和下面定义的 β_{kk}) 也可表示成几种形式, 恰如 β_{11} 有 (10.87) 和 (10.88) 这些等价的表示式一样(对二次函数). 后面, 我们将回到这个等价性来.

用类似的方法，也能建立

$$(\nabla f(x^i))^T \nabla f(x^j) = 0, \quad 0 \leq i < j \leq n, \tag{10.93}$$

$$\beta_{kj} = 0, \quad k \neq j, \tag{10.94}$$

$$\beta_k = \beta_{kk} = \frac{(\nabla f(x^k))^T \nabla f(x^k)}{(\nabla f(x^{k-1}))^T \nabla f(x^{k-1})}, \quad k = 1, \cdots, n. \tag{10.95}$$

既然当 $k \neq j$ 时 $\beta_{kj} = 0$, 在即将到来的讨论中, 我们就可去掉双重

下标 kk. 这样

$$z^{k+1}=-\nabla f(x^k)+\frac{(\nabla f(x^k))^T\nabla f(x^k)}{(\nabla f(x^{k-1}))^T\nabla f(x^{k-1})}z^k. \tag{10.96}$$

从这些推导得到的一个完整的无约束极小化算法称为**共轭梯度法**，它属于 Fletcher、Reeves[12]，其基础是 Hestenes、Stiefel[18] 为求解线性方程组提出的一个方法. 共轭梯度法可描述如下:

选择一个出发点 $x^0\in R^n$，求 $\nabla f(x^0)$ 的值，而后取 $z^1=-\nabla f(x^0)$. 依次沿着方向 $z^1, z^2, \cdots, z^n$ 极小化 $f(x)$ 而移动到 $x^1, x^2, \cdots, x^n$，其中 z^k 由(10.96)选定. 分别以 x^n 和 $-\nabla f(x^n)$ 作为新的 x^0 和 z^1，重新开始这一过程. 当下式满足时，算法结束:

$$(\nabla f(x^k))^T\nabla f(x^k)\leqslant\varepsilon, \tag{10.97}$$

其中 ε 是某个预先确定的小正数. 注意，若 f 由(10.62)给定，则方向 z^k 关于 Q 相互共轭. 这样，共轭梯度法具有二次终结性质.

例 10.2.1

用 Fletcher 和 Reeves 的共轭梯度法考察下面二次函数的极小化问题:

$$f(x)=\frac{3}{2}(x_1)^2+\frac{1}{2}(x_2)^2-x_1x_2-2x_1. \tag{10.98}$$

将 f 写成(10.62)的形式，便有

$$b=\begin{pmatrix}-2\\0\end{pmatrix},\quad Q=\begin{bmatrix}3&-1\\-1&1\end{bmatrix}. \tag{10.99}$$

假设从 $x^0=(-2,\ 4)^T$ 开始搜索，得到

$$\nabla f(x^0)=\begin{pmatrix}-12\\6\end{pmatrix},\quad z^1=\begin{pmatrix}12\\6\end{pmatrix}. \tag{10.100}$$

关于 α_1 极小化 $f(x^0+\alpha_1z^1)$，找到 $\alpha_1^*=\frac{5}{17}$. 于是

$$x^1=\begin{pmatrix}-2+\dfrac{(5)(12)}{17}\\[2ex]4-\dfrac{(5)(6)}{17}\end{pmatrix}=\begin{pmatrix}\dfrac{26}{17}\\[2ex]\dfrac{38}{17}\end{pmatrix},\quad \nabla f(x^1)=\begin{pmatrix}\dfrac{6}{17}\\[2ex]\dfrac{12}{17}\end{pmatrix}. \tag{10.101}$$

现在必须找第二个搜索方向 z^2, 由(10.96)有

$$z^2 = -\nabla f(x^1) + \frac{(\nabla f(x^1))^T \nabla f(x^1)}{(\nabla f(x^0))^T \nabla f(x^0)} z^1, \tag{10.102}$$

从而

$$z^2 = \begin{pmatrix} -\frac{6}{17} \\ -\frac{12}{17} \end{pmatrix} + \frac{\left(\frac{6}{17}\right)^2 + \left(\frac{12}{17}\right)^2}{(-12)^2 + (6)^2} \begin{pmatrix} 12 \\ -6 \end{pmatrix} = \begin{pmatrix} -\frac{90}{289} \\ -\frac{210}{289} \end{pmatrix}. \tag{10.103}$$

关于 α_2 极小化 $f(x^1 + \alpha_2 z^2)$, 得到 $\alpha_2^* = \frac{17}{10}$. 结果有

$$x^2 = \begin{pmatrix} \frac{26}{17} - \frac{(17)(90)}{(10)(289)} \\ \frac{38}{17} - \frac{(17)(210)}{(10)(289)} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tag{10.104}$$

象所断言的那样, 它是 f 的整体极小值点. 图 10.3 说明了这个算法的搜索方向和步骤. 】

要求读者证明, (10.95)中系数 β_k 至少存在三个等价的公式

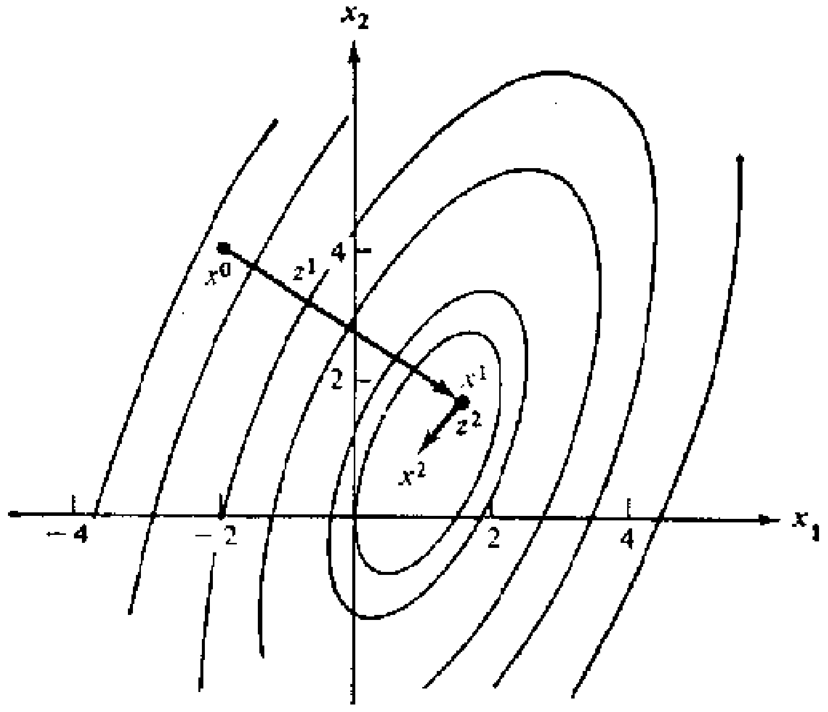


图 10.3 共轭梯度法

(当用于极小化二次函数时)．它们是

$$\beta_k=\frac{(z^k)^T\nabla^2 f(x^k)\nabla f(x^k)}{(z^k)^T\nabla^2 f(x^k)z^k},\tag{10.105}$$

$$\beta_k=\frac{(\gamma^k)^T\nabla f(x^k)}{(\gamma^k)^T z^k},\tag{10.106}$$

$$\beta_k=\frac{(\gamma^k)^T\nabla f(x^k)}{(\nabla f(x^{k-1}))^T\nabla f(x^{k-1})}.\tag{10.107}$$

这些公式在下述意义下等价：当用于极小化一个带有正定阵 Q 的二次函数，且初始方向取 $z^1=-\nabla f(x^0)$时，它们产生同样的搜索方向．公式(10.105)由Daniel[7]提出，公式(10.106)由Sorenson[34]和Wolfe[35]提出，而公式(10.107)用于Polak、Ribière[29]与Polyak[30]的共轭梯度算法中．Hestenes、Stiefel[18]发现，对于舍入误差，(10.106)不如(10.95)敏感．

初始方向 $z^1=-\nabla f(x^0)$的选取看来是十分重要的．例如，当我们选 $z^1\neq-\nabla f(x^0)$时，可能发生这样的情况：对于二次函数，Fletcher-Reeves 方法仅在重新开始的步骤以后才终结，而其他的形式在 n 步中达到极小值点．我们用简单的二次函数来说明上述断言[35]．

例 10.2.2

考虑下列函数的极小化：

$$f(x)=\frac{1}{2}(x_1)^2+\frac{1}{2}(x_2)^2.\tag{10.108}$$

设 $x^0=(1,1)$，不用负梯度方向 $-\nabla f(x^0)=(-1,-1)^T$，我们取 $z^1=(-1,0)^T$．沿 z^1 极小化，我们达到下一点 $x^1=(0,1)$，$\nabla f(x^1)=(0,1)$．利用(10.95)，下一个搜索方向为

$$z^2=-\begin{pmatrix}0\\1\end{pmatrix}+\frac{1}{2}\begin{pmatrix}-1\\0\end{pmatrix}=\begin{pmatrix}-\frac{1}{2}\\-1\end{pmatrix},\tag{10.109}$$

$x^2=\left(-\frac{2}{5},\ \frac{1}{5}\right)$．不重新开始的话，我们将在一条趋近于原点的螺旋形轨线上移动．另一方面，用(10.106)或(10.107)，因为 $(\gamma^1)^T\nabla f(x^1)=0$，我们得到 $z^2=(0,-1)^T$，z^2 方向上的极小值在点

$x^2=(0, 0)$达到，当然 $\nabla f(x^2)=0$，所以算法终结. ∎

上述例子中，用式(10.95)不能在两步中求得精确极小值点，其理由是，仅当以前的步骤是准确的并且特别取 $z^1=-\nabla f(x^0)$时，这公式才能保证在第 k 次迭代中有 $(z^k)^T Q z^{k+1}=0$. 而先前给出的其他公式即使 $z^1\neq-\nabla f(x^0)$，也能保证相继的两个搜索方向 z^k 和 z^{k+1} 关于 Q 是共轭的；但是在这情形下，对 $n>2$，一般不能得到 n 个搜索方向的相互共轭性，因此二次终结可能不出现. 由此可推知，当用共轭梯度法来求二次函数的极小时，如不作"标准的"选取 $z^1=-\nabla f(x^0)$ 和不作周期性的重新开始，将使其效率大为降低. 后面我们还将回到这一点上来.

10.3 共轭梯度法的收敛性

关于收敛性问题，至今我们仅知道，正常执行的共轭梯度法具有二次终结性质. 现在我们证明共轭梯度法对于一大类函数的趋于整体极小的收敛性. 共轭梯度法的不同形式的收敛性质在文献中已有广泛讨论，例如见 Daniel[10]，Elkin[11]，McCormick、Ritter[21, 22]，Ortega、Rheinboldt[25]，Polak[28]，Polak、Ribière[29] 和 Polyak[30]. 这里，我们证明不作重新开始的 Polak-Ribière-Polyak 共轭梯度法(简记为 PRP 方法)的收敛性. PRP 算法可描述如下：设 $x^0\in R^n$ 已给定，选 z^1 使得

$$(z^1)^T\nabla f(x^0)<0, \tag{10.110}$$

对 $k=1, 2, \cdots$，令

$$z^{k+1}=-\nabla f(x^k)+\beta_k z^k, \tag{10.111}$$

其中 β_k 由(10.107)给出，且

$$x^k=x^{k-1}+\alpha_k^* z^k, \tag{10.112}$$

其中 α_k^* 极小化 $f(x^{k-1}+\alpha_k z^k)$，即 α_k^* 是用精确的线性搜索求得的. 我们的证明主要基于 Daniel[10] 和 McCormick、Ritter[22].

我们从证明 PRP 算法是严格下降算法开始.

引理 10.4

设 f 是 R^n 中一个开子集上的连续可微实函数，则对于由

PRP 算法产生的点列$\{x^k\}$，每当$\nabla f(x^k)\neq 0$时总有

$$f(x^k)>f(x^{k+1}),\quad k=0,\ 1,\ \cdots. \tag{10.113}$$

【证明】 用$\nabla f(x^k)\neq 0$乘(10.111)，我们得到

$$(\nabla f(x^k))^T z^{k+1}=-(\nabla f(x^k))^T\nabla f(x^k)<0,\quad k=1,\ 2,\ \cdots. \tag{10.114}$$

这是因为由α_k^*的选取，$(\nabla f(x^k))^T z^k=0$. 所以，对充分小的α_{k+1}，从(10.110)和(10.114)推得

$$f(x^k)>f(x^k+\alpha_{k+1}z^{k+1}). \tag{10.115}$$

因为$x^{k+1}=x^k+\alpha_{k+1}^*z^{k+1}$，其中$\alpha_{k+1}^*$极小化$f(x^k+\alpha_{k+1}z^k)$，所以不等式(10.113)成立. 】

下一个定理给出的关于 PRP 方法收敛性的证明中，我们假定f是在某开凸集$C\subset R^n$上的二次连续可微实函数，Hesse 阵$\nabla^2 f$是正定的（当然这蕴涵f是严格凸的）. 还假设，对每个$x^0\in C$，$f(x^0)=\alpha$，由式

$$S(f,\ \alpha)=\{x: x\in C,\ f(x)\leqslant\alpha\} \tag{10.116}$$

定义的水平集$S(f,\ \alpha)\subset C$是紧致的. 水平集的紧致性是在极小化算法的收敛性定理中广泛使用的一个假定. 因此，考虑关于函数ϕ的某些条件以确保其水平集为紧致，是适宜的. 注意，对于连续函数ϕ，所有水平集是闭的，因此这些集为紧致的充要条件是它们有界. 所有水平集的紧致性可由下面定理刻划.

定理 10.5

设ϕ是无界集$T\subset R^n$上的连续实函数，则对于任意$\alpha\in R$，非空水平集

$$S(\phi,\ \alpha)=\{x: x\in T,\ \phi(x)\leqslant\alpha\} \tag{10.117}$$

紧致的充要条件是：

$$\lim_{k\to\infty}\{\phi(x^k)\}=+\infty \tag{10.118}$$

对于满足$x^k\in T$且

$$\lim_{k\to\infty}\{\|x^k\|\}=+\infty \tag{10.119}$$

的每个序列$\{x^k\}$都成立.

【证明】 假定ϕ的所有水平集有界，令$\{x^k\}$是满足(10.119)的T中点列，并设存在实数β，使得对无穷多个k，$\phi(x^k)\leqslant\beta$. 则存在$\{x^k\}$的子序列$\{x^{k_m}\}$，使得对每个x^{k_m}有$\phi(x^{k_m})\leqslant\beta$. 这样，$\{x^{k_m}\}$的所有点在有界集$S(\phi, \beta)$中，这与(10.119)矛盾. 反之，假设(10.118)与(10.119)成立，而对某$\hat{\alpha}$，$S(\phi, \hat{\alpha})$无界，则存在序列$\{x^k\}$，$x^k\in S(\phi, \hat{\alpha})$且满足(10.119)和$\phi(x^k)\leqslant\hat{\alpha}$，这与(10.118)矛盾. 】

现在假设f是C上的二次连续可微函数，且存在正数μ与η，使得对每个$x\in C$和$y\in R^n$，我们有

$$\mu\|y\|^2\leqslant y^T\nabla^2 f(x)y\leqslant\eta\|y\|^2. \tag{10.120}$$

或者等价地，$\nabla^2 f(x)$的特征值介于μ和η之间. 这些不等式也明显地蕴涵着，对每个$x\in C$，Hesse 阵$\nabla^2 f(x)$是正定的，所以f在C上是严格凸的. 我们来证明 PRP 方法的搜索方向是有界的.

引理 10.6

假设(10.120)成立，且z^k，x^k，α_k由 PRP 方法确定，则

$$\|\nabla f(x^k)\|\leqslant\|z^{k+1}\|\leqslant\frac{\mu+\eta}{\mu}\|\nabla f(x^k)\|,\quad k=1, 2, \cdots. \tag{10.121}$$

【证明】 我们有

$$(\nabla f(x^k))^T z^k=0,\quad k=1, 2, \cdots, \tag{10.122}$$

$$(\nabla f(x^k))^T z^{k+1}=-\|\nabla f(x^k)\|^2,\quad k=1, 2, \cdots, \tag{10.123}$$

这样β_k能写成

$$\beta_k=\frac{(\gamma^k)^T\nabla f(x^k)}{(\gamma^k)^T z^k}. \tag{10.124}$$

由中值定理，有

$$\nabla f(x^k)=\nabla f(x^{k-1})+\nabla^2 f(\xi^{k-1})(x^k-x^{k-1}) \tag{10.125}$$

$$=\nabla f(x^{k-1})+\nabla^2 f(\xi^{k-1})\alpha_k^* z^k, \tag{10.126}$$

其中ξ^{k-1}是x^k和x^{k-1}连线上某一点. 因此得到

$$\gamma^k=\nabla^2 f(\xi^{k-1})\alpha_k^* z^k, \tag{10.127}$$

$$|(\gamma^k)^T\nabla f(x^k)|=|\alpha_k^*(z^k)^T\nabla^2 f(\xi^{k-1})\nabla f(x^k)|. \tag{10.128}$$

由 Cauchy-Schwarz 不等式[2]，有

$$|(\gamma^k)^T\nabla f(x^k)| \leqslant \alpha_k^* \|\nabla f(x^k)\| \|\nabla^2 f(\xi^{k-1})z^k\| \tag{10.129}$$

$$\leqslant \alpha_k^* \|\nabla f(x^k)\| \left\|\nabla^2 f(\xi^{k-1})\frac{z^k}{\|z^k\|}\right\| \|z^k\| \tag{10.130}$$

$$\leqslant \alpha_k^* \|\nabla f(x^k)\| \|z^k\| \{\sup_{\|z\|=1} \|\nabla^2 f(\xi^{k-1})z^k\|\} \tag{10.131}$$

$$\leqslant \alpha_k^* \|\nabla f(x^k)\| \|z^k\| \|\nabla^2 f(\xi^{k-1})\| \tag{10.132}$$

$$\leqslant \alpha_k^* \|\nabla f(x^k)\| \|z^k\| \eta, \tag{10.133}$$

其中 (10.132) 和 (10.133) 由对称阵的模 $\|A\|$ 的定义[15] 与 (10.120)推得. 如第 1 章中注意到的, 成立

$$\|A\| = \sup_{\|x\|=1} \|Ax\| = \sup_{\|x\|=1} x^T A x. \tag{10.134}$$

类似地, 可以证明

$$|(\gamma^k)^T z^k| \geqslant \alpha_k^* \mu \|z^k\|^2. \tag{10.135}$$

于是就有

$$|\beta_k| \leqslant \frac{\eta\|\nabla f(x^k)\|}{\mu\|z^k\|}, \tag{10.136}$$

$$\|\beta_k z^k\| \leqslant \frac{\eta}{\mu}\|\nabla f(x^k)\|. \tag{10.137}$$

用(10.111), (10.137)和三角不等式[2], 我们得到

$$\|z^{k+1}\| \leqslant \|\nabla f(x^k)\| + \|\beta_k z^k\| \leqslant \frac{\mu+\eta}{\mu}\|\nabla f(x^k)\|. \tag{10.138}$$

现在, 由

$$\begin{aligned}\|z^{k+1}\|^2 = &-(\nabla f(x^k))^T[-\nabla f(x^k) + \beta_k z^k] \\ &+ \beta_k (z^k)^T[-\nabla f(x^k) + \beta_k z^k],\end{aligned} \tag{10.139}$$

且由(10.122)可推得

$$\|z^{k+1}\|^2 = \|\nabla f(x^k)\|^2 + (\beta_k)^2 \|z^k\|^2 \geqslant \|\nabla f(x^k)\|^2, \tag{10.140}$$

所以

$$\|z^{k+1}\| \geqslant \|\nabla f(x^k)\|. \tag{10.141}$$

∎

我们现在叙述并证明 PRP 方法的下述收敛性定理.

定理 10.7

设 f 是开凸集 $C \subset R^n$ 上的二次连续可微实函数, 且满足

(10.120). 令 $x^0 \in C$, $f(x^0)=\alpha$, 并设 $S(f, \alpha)$是紧致的. 则 PRP 算法是严格下降方法, 即每当 $\nabla f(x^k)\neq 0$ 时, 总有

$$f(x^k)>f(x^{k+1}), \quad k=0, 1, \cdots. \tag{10.142}$$

并且, 或者对某个 k 有 $\nabla f(x^k)=0$; 或者有

$$\lim_{k\to\infty}\{\|\nabla f(x^k)\|\}=0. \tag{10.143}$$

序列$\{x^k\}$收敛到点 $x^*\in S(f, \alpha)$, 在这点处 $\nabla f(x^*)=0$, 从而 x^* 是 f 的严格整体极小值点.

【证明】 由引理 10.4, PRP 算法是严格下降方法, 从而(10.142)成立. 由(10.127), 我们有

$$(\gamma^{k+1})^T z^{k+1}=\alpha_{k+1}^*(z^{k+1})^T\nabla^2 f(\xi^k)z^{k+1}, \tag{10.144}$$

于是

$$\alpha_{k+1}^*=\frac{-(\nabla f(x^k))^T z^{k+1}}{(z^{k+1})^T\nabla^2 f(\xi^k)z^{k+1}}=\frac{\|\nabla f(x^k)\|^2}{(z^{k+1})^T\nabla^2 f(\xi^k)z^{k+1}}. \tag{10.145}$$

从(10.120)和引理 10.6 我们得到

$$\alpha_{k+1}^*\geqslant\frac{\|\nabla f(x^k)\|^2}{\eta\|z^{k+1}\|^2}\geqslant\frac{(\mu)^2}{\eta(\mu+\eta)^2}. \tag{10.146}$$

对所有 $0<\alpha_{k+1}\leqslant\alpha_{k+1}^*$, 用 Taylor 定理, 我们能写出

$$f(x^k+\alpha_{k+1}z^{k+1})=f(x^k)+\alpha_{k+1}(\nabla f(x^k))^T z^{k+1} + \frac{1}{2}(\alpha_{k+1})^2(z^{k+1})^T\nabla^2 f(\hat{\xi}^k)z^{k+1}, \tag{10.147}$$

其中 $\hat{\xi}^k$ 是 $x^k+\alpha_{k+1}z^{k+1}$ 和 x^k 连线上某一点. 由(10.120), (10.121)和(10.123), 有

$$f(x^k+\alpha_{k+1}z^{k+1})\leqslant f(x^k)-\alpha_{k+1}\|\nabla f(x^k)\|^2 + \frac{1}{2}(\alpha_{k+1})^2\eta\frac{(\mu+\eta)^2}{(\mu)^2}\|\nabla f(x^k)\|^2. \tag{10.148}$$

令

$$\alpha_{k+1}=\frac{(\mu)^2}{\eta(\mu+\eta)^2}, \tag{10.149}$$

并按 α_{k+1}^* 和 x^{k+1} 的定义, 我们得到

$$f(x^{k+1}) \leqslant f\left(x^k + \frac{(\mu)^2}{\eta(\mu+\eta)^2} z^{k+1}\right) \tag{10.150}$$

$$\leqslant f(x^k) - \frac{(\mu)^2}{\eta(\mu+\eta)^2} \|\nabla f(x^k)\|^2$$

$$+ \frac{1}{2} \frac{(\mu)^2}{\eta(\mu+\eta)^2} \|\nabla f(x^k)\|^2 \tag{10.151}$$

$$\leqslant f(x^k) - \frac{1}{2} \frac{(\mu)^2}{\eta(\mu+\eta)^2} \|\nabla f(x^k)\|^2, \tag{10.152}$$

于是

$$f(x^k) - f(x^{k+1}) \geqslant \frac{1}{2} \frac{(\mu)^2}{\eta(\mu+\eta)^2} \|\nabla f(x^k)\|^2. \tag{10.153}$$

因为序列 $\{x^k\}$ 收敛到紧致集 $S(f, \alpha)$ 中的点 x^*，而函数 f 在 $S(f, \alpha)$ 上有下界，从 (10.153) 推知 $\{\|\nabla f(x^k)\|\}$ 必收敛于零*，所以 $\nabla f(x^*)=0$. 因为 f 是严格凸的，所以点 x^* 满足第 4 章给出的、关于 f 的严格整体极小值点的充分条件. ▮

类似于 Fletcher-Reeves 方法，我们可以令

$$z^{k+1} = -\nabla f(x^k), \quad k=p,\ 2p,\ 3p,\ \cdots \tag{10.154}$$

来定义带有周期性重新开始的 PRP 方法，其中 p 通常选为 n 或 $n+1$，这过程的其余部分如同早先所规定的. 读者能证明，对带有周期性重新开始的 PRP 方法，定理 10.7 也成立. 对于 PRP 方法的这两种形式的较弱的结果可叙述如下：

定理 10.8

设 f 是开凸集 $C \subset R^n$ 上的连续可微实函数，令 $x^0 \in C$，$f(x^0) = \alpha$，并设 $S(f, \alpha)$ 紧致. 则带有或不带有周期性重新开始的 PRP 算法是严格下降方法；就是说，若 $\nabla f(x^k) \neq 0$，便有

$$f(x^k) > f(x^{k+1}), \quad k=0,\ 1,\ \cdots. \tag{10.155}$$

并且，或者对某个 k，$\nabla f(x^k)=0$；或者有

$$\lim_{k\to\infty} \{\|\nabla f(x^k)\|\} = 0. \tag{10.156}$$

序列 $\{x^k\}$ 的每个凝聚点是 $\nabla f(x)=0$ 的解；并且，如果它有唯一解

---

* 译注：由 f 在 $S(f, \alpha)$ 上有下界及 (10.153)，就可说明凝聚点必为逗留点，再利用凸函数的性质可说明凝聚点的唯一性，于是 $\{x^k\}$ 收敛.

$x^* \in S(f, \alpha)$, 则 x^* 是 f 在 C 上的整体极小值点.

证明与前面所述相类似, 从略.

现在转到共轭梯度法的收敛速率. 正如我们已知道的, 如果初始方向 z^1 选为 $-\nabla f(x^0)$, 则这些方法有二次终结性质. Crowder、Wolfe[5] 证明了, 一般来说, 一些共轭梯度法的收敛速率不坏于最速下降法; 即它不坏于线性收敛. 更特殊地, 令 $f(x) = \frac{1}{2}x^TQx, x \in R^n$, 其中 Q 是对称正定阵. 设在第 k 次迭代时已有点 x^{k-1} 和方向 z^k, 关于 α_k 极小化 $f(x^{k-1}+\alpha_k z^k)$, 我们达到点 x^k, 它满足 $(z^k)^T\nabla f(x^k)=0$. 现在, 我们可以如最速下降法那样选 $z^{k+1} = -\nabla f(x^k)$, 或者如 PRP 方法那样, 用共轭梯度公式

$$z^{k+1} = -\nabla f(x^k) + \beta_k z^k, \tag{10.157}$$

其中 β_k 由 PRP 方法中的(10.105)、(10.106)或(10.107)给出. 接着我们关于 α_{k+1} 极小化 $f(x^k+\alpha_{k+1}z^{k+1})$, 并达到一点 x^{k+1}. 方向 z^{k+1} 的两种不同形式一般将导致不同的两点 x^{k+1}. 在最速下降(SD)移动的情形, 我们得到

$$f(x_{SD}^{k+1}) = \frac{1}{2}(\nabla f(x^k))^TQ^{-1}\nabla f(x^k) - \frac{1}{2}\frac{\|\nabla f(x^k)\|^4}{(\nabla f(x^k))^TQ\nabla f(x^k)}, \tag{10.158}$$

而在共轭梯度(CG)移动的情形, 我们得到

$$(z^{k+1})^TQz^{k+1} = (\nabla f(x^k))^TQ\nabla f(x^k) - \frac{[(z^k)^TQ\nabla f(x^k)]^2}{(z^k)^TQz^k}, \tag{10.159}$$

$$f(x_{CG}^{k+1}) = \frac{1}{2}(\nabla f(x^k))^TQ^{-1}\nabla f(x^k) - \frac{1}{2}\frac{\|\nabla f(x^k)\|^4}{(z^{k+1})^TQz^{k+1}}. \tag{10.160}$$

所以由(10.158)至(10.160), 有

$$f(x_{CG}^{k+1}) \leqslant f(x_{SD}^{k+1}). \tag{10.161}$$

换言之, 任一共轭梯度法的每次迭代中 f 所减小的值至少和最速下降法一样. 既然最速下降法有线性收敛速率, 我们就可断定, β_k 由(10.105)至(10.107)定义的共轭梯度法有不坏于线性速率的收

敛速率．然而，例 10.2.2 表明，用(10.95)选 β_k 时，不等式(10.161)不一定成立，这是因为用 Fletcher-Reeves 方法选取的方向 z^1 和 z^2 不共轭．Crowder 和 Wolfe[5] 也证明了，不用"标准的"初始方向 $z^1=-\nabla f(x^0)$，共轭梯度法的收敛速率可能象线性速率那样慢．如我们将看到的，他们用一个简单的二次函数，说明这种共轭梯度法可以只有线性收敛速率．

例 10.3.1

令 $f(x)=\dfrac{1}{2}x^TQx$，其中

$$Q=\begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{10.162}$$

然后假设我们从点 $x^0=(10/\sqrt{6},\ -\sqrt{5}/\sqrt{6},\ 0)$ 开始搜索，且选 $z^1=(-5/2\sqrt{6},\ 7/2\sqrt{30},\ -3/4\sqrt{5})^T$．注意 $\nabla f(x^k)=Qx^k$ 和

$$(z^1)^T\nabla f(x^0)=\frac{-5}{2\sqrt{6}}\ \frac{1}{\sqrt{6}}+\frac{7}{2\sqrt{30}}\ \frac{(-5)}{\sqrt{6}}<0, \tag{10.163}$$

z^1 是一个下降方向．其次，设 z^{k+1} 由(10.157)给出，其中 β_k 由公式(10.105)选取．可以证明对所有 $k=1,2,\cdots$，

$$\alpha_k^*=\frac{8}{5},\qquad \beta_k=\frac{9}{25}. \tag{10.164}$$

进一步有

$$\nabla f(x^{k+1})=rR\nabla f(x^k),\qquad z^{k+1}=rRz^k, \tag{10.165}$$

其中 $r=\dfrac{3}{5}$，且

$$R=\begin{bmatrix} 1 & 0 & 0 \\ 0 & -\dfrac{1}{5} & -\dfrac{2\sqrt{6}}{5} \\ 0 & -\dfrac{2\sqrt{6}}{5} & -\dfrac{1}{5} \end{bmatrix} \tag{10.166}$$

是正交阵．因此

$$\frac{\|x^{k+1}\|}{\|x^k\|}=\frac{3}{5},\quad \frac{\|f(x^{k+1})\|}{\|f(x^k)\|}=\frac{9}{25} \tag{10.167}$$

从而收敛速率是线性的. ∎

利用初始搜索方向 $z^1=-\nabla f(x^0)$ 和周期性的重新开始, Cohen[4], Daniel[10], McCormick、Ritter[22], Ortega、Rheinboldt[25], Polak[28] 和 Polyak[30] 证明了, 在如(10.120)那样的关于 f 的适当假定下, 各种形式的共轭梯度法的收敛速率是"n 步超线性"的, 即

$$\lim_{k\to\infty}\frac{\|x^{k+n}-x^*\|}{\|x^k-x^*\|}=0, \tag{10.168}$$

其中 x^* 是 f 的极小值点. 其中有几篇著作还证明了, 实际的速率是 n 步二次的. Daniel[7,8,9] 证明, 对于他那个利用(10.105)并带有"标准"初始方向 $z^1=-\nabla f(x^0)$, 但不带有周期性重新开始的共轭梯度法, 对充分大的 k 有

$$\|x^{k+n}-x^*\|\leqslant q\|x^k-x^*\|^{[1+1/(4n-3)]}. \tag{10.169}$$

根据理论的收敛速率, 我们可以断定, 宁可采用带有周期性重新开始的共轭梯度法, 而不采用连续的形式. 共轭梯度法以及某些 Newton 型方法如 Goldstein-Price 方法或 Ritter 方法, 它们的收敛速率的一个有趣的讨论可以在 McCormick、Ritter[21]中找到. 看来, 某些算法是否受欢迎一般不仅仅取决于收敛速率, 而且还要考虑其他方面, 如每次迭代所需的计算量, 需用的计算机内存容量等. 例如, 应当指出, 共轭梯度法的收敛性强烈地依赖于精确的一维最优化, 而这是代价很大的. 不过, Klessig、Polak[20] 修改了 PRP 方法, 免去了一维搜索, 但保留了某些共轭性质.

Ortega、Rheinboldt[24,25,26] 对于本章所述算法的收敛性质采用统一的处理. 他们考虑 R^n 上的连续可微实函数 f 和由

$$x^k=x^{k-1}+\alpha_k z^k,\quad z^k\neq 0 \tag{10.170}$$

定义的迭代下降方法. 为了保证这种方法的收敛性, 步长 α_k 将选得在每次迭代中 f"充分"减小. 这概念是借助强迫函数 σ 来定义的, σ 的定义域和值域是非负数的集合且具有性质: 对非负数的任

意序列$\{t^k\}$,

$$\lim_{k\to\infty}\{\sigma(t^k)\}=0 \quad 蕴涵 \quad \lim_{k\to\infty}\{t^k\}=0. \tag{10.171}$$

例如$\sigma(t)=ct$,其中$c>0$,就满足上述关系式. 如果存在不依赖于x^k的函数$\hat{\sigma}$,使得下式成立:

$$f(x^{k-1})-f(x^k)\geqslant\hat{\sigma}\frac{|(z^k)^T\nabla f(x^{k-1})|}{\|z^k\|}, \tag{10.172}$$

就说迭代下降方法使f在任意点x^k处有充分减小. 如果f有下界,则当$k\to\infty$时,$f(x^{k-1})-f(x^k)$趋于零,且

$$\lim_{k\to\infty}\left\{\frac{(z^k)^T\nabla f(x^{k-1})}{\|z^k\|}\right\}=0. \tag{10.173}$$

现在我们希望说,式(10.173)蕴涵

$$\lim_{k\to\infty}\{\nabla f(x^k)\}=0. \tag{10.174}$$

这断言依赖于z^k的选择. Ortega 和 Rheinboldt 定义R^n中非零向量序列$\{z^k\}$为与序列$\{x^k\}$, $x^k\in R^n$, **梯度相关的**, 如果对某个强迫函数σ有

$$\frac{|(z^k)^T\nabla f(x^{k-1})|}{\|z^k\|}\geqslant\sigma(\|\nabla f(x^{k-1})\|). \tag{10.175}$$

例如,若$z^k=-\nabla f(x^k)$, 则上述关系式关于$\sigma(t)=t$成立. 因此推知, 如果由下降方法产生的点列$\{x^k\}$对于与之梯度相关的序列$\{z^k\}$满足(10.173),则(10.175)蕴涵(10.174), 从而得到了趋于f的逗留点的收敛性. 特别地,由最速下降法、某种限步 Newton 法和共轭梯度法选取的方向已被证明是梯度相关的. 这样, 类似于本章所给出的收敛性结果可以换用这种方式导出. 收敛性结果的一个有趣的拓广是, 用梯度相关向量作周期性重新开始的下降方法也是收敛的,不管中间的下降方向怎样选取.

最后要注意的是, 共轭梯度法的很多背景至今尚未完全理解. 如周期性的重新开始, 初始搜索方向的选取或者回避精确的线性搜索等怎样影响共轭梯度法的执行, 这样一些问题仍有待回答.

练　习

10.A. 对于一切无约束最优化算法的一个"经典的"试验函数，是 Rosenbrock 的二变量"弯谷"函数[32]：

$$f(x)=100[x_2-(x_1)^2]^2+(1-x_1)^2, \tag{10.176}$$

它有唯一的极小值点 $x^*=(1, 1)$.

(a) 用(10.5)给出的式子计算 f 在点$(-1.2, 1)$, $(0, 0)$和$(1.5, 1)$的二次近似，并求它们的极小值点借此观察 Newton 法用于这函数时的性态. 比较在二次函数所预测的极小值点上 f 的值与在所给的点上 f 的值. 在每种情况下函数值是否都减小了?

(b) 求在前面三点的每一点上 f 的梯度之值，比较最速下降方向与指向 f 之极小值点的方向.

10.B. 对于(9.194)所给的函数，从 $x^0=(-4, 6)^T$ 出发的 Newton 方向是什么? 计算用 Newton 法到达的相继两点.

10.C. 证明，(10.18) 和 (10.19) 提出的求最速下降方向的问题是一个凸规划问题，其解由式(10.20)给出. 对于以(10.23)代替 (10.19) 的情形，证明同样的结论，且解由(10.24)给出.

10.D. 将最速下降法用于练习 9.A 中的二次函数(9.188)和(9.189)，取它们各自的出发点 x^0，确定其相对效率 η，η 由(10.27)给出. 对这两种情形计算(10.29)给出的 η 的下界.

10.E. 假设我们用下式修改最速下降法：

$$x^{k+1}=x^k-\theta_k^* B(x^k)\nabla f(x^k), \tag{10.177}$$

其中 $B(x)$ 是 $n\times n$ 正定阵，其元素是 x 的连续函数，θ_k^* 是由极小化 $f(x^k-\theta_k B(x^k)\nabla f(x^k))$ 所确定的数量. 证明对于这方法的与定理 10.1 相似的定理.

10.F. 画出 Goldstein-Price 算法的框图，用这算法去求(9.194)给出的函数的极小值点，从 $(-4, 6)^T$ 出发. 执行三次迭代并比较所得到的 $Q(x^k)$ 和实际的 Hesse 阵 $\nabla^2 f(x^k)$.

10.G. 证明方程(10.93)至(10.95).

10.H. 极小化 R^n 上可微函数的一个下降算法构造如下，从中导出公式(10.106)：选一点 $x^0\in R^n$ 和搜索方向 $s^1=-\nabla f(x^0)$. 沿 s^1 作线性搜索达到 x^1，即$(s^1)^T\nabla f(x^1)=0$. 令 s^2 满足下述要求：

(a) 它正交于 γ^1,

(b) 它是下降方向，即 $(s^2)^T\nabla f(x^1)<0$.

(c) 它是 s^1 和 $\nabla f(x^1)$ 的线性组合.

证明 s^2 能写成

$$s^2=-\nabla f(x^1)+\beta_1 s^1, \tag{10.178}$$

其中

$$\beta_1=\frac{(\gamma^1)^T\nabla f(x^1)}{(\gamma^1)^T s^1}. \tag{10.179}$$

如果后面所有 s^k 以类似方式定义，证明相应的 β_k 由(10.106)给出. 注意这些推导并未假定 f 是二次的.

10.I. 证明，公式(10.95)，(10.105)，(10.106)和(10.107)用于极小化一个带正定阵 Q 的二次函数时，产生相同的搜索方向. 对于一般函数，在什么条件下，(10.106)和(10.107)将产生相同的方向?

10.J. 对于(9.194)给出的函数，用 Fletcher-Reeves 算法进行三次迭代，从 $(-4, 6)^T$ 出发，并使每次线性搜索是第 8 章所描述的单个二次近似. 将所达到的点与练习 10.F 中所得到的点进行比较.

10.K. Shah，Buehler 和 Kempthorne[83] 提出了极小化 R^n 上可微函数的下述算法，称为**平行切线法**或简称 **Partan 方法**. 给定 x^0，令 $s^1=-\nabla f(x^0)$，沿 s^1 极小化 f 达到点 x^1. 令 $s^2=-\nabla f(x^1)$，关于 α_2 极小化 $f(x^1+\alpha_2 s^2)$，取 $x^2=x^1+\alpha_2^* s^2$，再令 $s^3=x^2-x^0$，这称为加速方向. 沿 s^3 极小化 f 达到 x^3. 达到 x^3 之后，交替地沿负梯度方向和加速方向相继作线性搜索，即令

$$s^k=-\nabla f(x^{k-1}), \qquad k=1, 2, 4, 6, \cdots, \tag{10.180}$$

$$s^k=(x^{k-1}-x^{k-3}), \quad k=3, 5, 7, 9, \cdots. \tag{10.181}$$

对于例 10.2.1 的二次函数，从 $(-2, 4)^T$ 出发试试这方法，至少走四步.

10.L. 证明 Partan 方法用于 R^n 上带正定阵 Q 的二次函数时有下述性质:

(a) 向量 (x^1-x^0)，(x^3-x^1)，$\cdots$，$(x^{2n-1}-x^{2n-3})$ 关于 Q 相互共轭.

(b) 点 x^3 是在由向量 s^1 和 $\nabla f(x^1)$ 所生成的仿射集上 f 的极小值点. 类似地，x^5 是 s^1、$\nabla f(x^1)$ 和 $\nabla f(x^3)$ 所生成的仿射集上 f 的极小值点，$\cdots$，x^{2n-1} 是 s^1、$\nabla f(x^1)$、$\nabla f(x^3)$、$\cdots$ 和 $\nabla f(x^{2n-3})$ 所生成的仿射集上 f 的极小值点.

(c) 向量 $\nabla f(x^0)$，$\nabla f(x^1)$，$\nabla f(x^3)$，$\cdots$，$\nabla f(x^{2n-3})$ 相互正交.

(d) Partan 方法具有二次终结性质.

10.M. 导出方程(10.158)至(10.161).

10.N. 建立例 10.3.1 中的关系式(10.164)至(10.167).

参 考 文 献

1. Akaike, H., "On a Successive Transformation of Probability Distribution and its Application to the Analysis of the Optimum Gradient Method," *Ann. Inst. Stat. Math. Tokyo*, **11**, 1–16 (1959).
2. Bartle, R. G., *The Elements of Real Analysis*, 2nd ed., John Wiley & Sons, New York, 1976.
3. Cauchy, A., "Méthode Générale pour la Résolution des Systéms d'Équations Simultanées," *Comp. Rend. Acad. Sci. Paris*, pp. 536–538 (1847).
4. Cohen, A., "Rate of Convergence for Root Finding and Optimization Algorithms," *SIAM J. Numer. Anal.*, **9**, 248–259 (1972).
5. Crowder, H., and P. Wolfe, "Linear Convergence of the Conjugate Gradient Method," *IBM J. Res. & Dev.*, **16**, 431–433 (1972).
6. Curry, H. B., "The Method of Steepest Descent for Non-Linear Minimization Problems," *Quart. Appl. Math.*, **2**, 258–261 (1944).
7. Daniel, J. W., "The Conjugate Gradient Method for Linear and Nonlinear Operator Equations," *SIAM J. Numer. Anal.*, **4**, 10–26 (1967).
8. Daniel, J. W., "Convergence of the Conjugate Gradient Method with Computationally Convenient Modifications," *Numerische Mathematik*, **10**, 125–131 (1967).
9. Daniel, J. W., "A Correction Concerning the Convergence Rate for the Conjugate Gradient Method," *SIAM J. Numer. Anal.*, **7**, 277–280 (1970).
10. Daniel, J. W., *The Approximate Minimization of Functionals*, Prentice-Hall, Englewood Cliffs, N.J., 1971.
11. Elkin, R., "Convergence Theorems for Gauss-Seidel and Other Minimization Algorithms," Ph.D. dissertation, University of Maryland, College Park, 1968.
12. Fletcher, R., and C. M. Reeves, "Function Minimization by Conjugate Gradients," *Computer J.*, **7**, 149–154 (1964).
13. Forsythe, G. E., "On the Asymptotic Directions of the s-Dimensional Optimum Gradient Method," *Numerische Mathematik*, **11**, 57–76 (1968).
14. Goldfeld, S. M., R. E. Quandt, and H. F. Trotter, "Maximization by Quadratic Hill-Climbing," *Econometrica*, **34**, 541–551, (1966).
15. Goldstein, A. A., *Constructive Real Analysis*, Harper & Row, New York, 1967.
16. Goldstein, A. A., and J. F. Price, "An Effective Algorithm for Minimization," *Numerische Mathematik*, **10**, 184–189 (1967).
17. Greenstadt, J., "On the Relative Efficiencies of Gradient Methods," *Math. Comp.*, **21**, 360–367 (1967).

18. Hestenes, M. R., and E. Stiefel, "Methods of Conjugate Gradients for Solving Linear Systems," *J. Res. Nat. Bur. Stand.*, 49, 409–436 (1952).

19. Kantorovich, L. V., and G. P. Akilov, *Functional Analysis in Normed Spaces*, Pergamon Press, Oxford, 1964.

20. Klessig, R., and E. Polak, "Efficient Implementations of the Polak-Ribière Conjugate Gradient Algorithm," *SIAM J. Control*, 10, 524–549 (1972).

21. McCormick, G. P., and K. Ritter, "Methods of Conjugate Directions Versus Quasi-Newton Methods," *Math. Prog.*, 3, 101–116 (1972).

22. McCormick, G. P., and K. Ritter, "Alternative Proofs of the Convergence Properties of the Conjugate Gradient Method," *J. Optimization Theory & Appl.*, 13, 497–518 (1974).

23. Murray, W., "Second Derivative Methods," in *Numerical Methods of Unconstrained Optimization*, W. Murray (Ed.), Academic Press, London, 1972.

24. Ortega, J. M., and W. C. Rheinboldt, "Local and Global Convergence of Generalized Linear Iterations," in *Numerical Solution of Nonlinear Problems*, J. M. Ortega, and W. C. Rheinboldt (Eds.), SIAM, Philadelphia, 1970.

25. Ortega, J. M., and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

26. Ortega, J. M., and W. C. Rheinboldt, "A General Convergence Result for Unconstrained Minimization Methods," *SIAM J. Numer. Anal.*, 9, 40–43 (1972).

27. Ostrowski, A. M., *Solution of Equations and Systems of Equations*, 2nd ed., Academic Press, New York, 1966.

28. Polak, E., *Computational Methods in Optimization*, Academic Press, New York, 1971.

29. Polak, E., and G. Ribière, "Note sur la Convergence de Méthodes de Directions Conjugées," *Rev. Fr. Inform. Rech. Oper.*, 16-R1, 35–43 (1969).

30. Polyak, B. T., "The Conjugate Gradient Method in Extremal Problems," *USSR Comp. Maths. and Math. Phys.*, 9, No. 4, 94–112 (1969).

31. Ritter, K., "A Superlinearly Convergent Method for Unconstrained Minimization," in *Nonlinear Programming*, J. B. Rosen, O. L. Mangasarian, and K. Ritter (Eds.), Academic Press, New York, 1970.

32. Rosenbrock, H. H., "An Automatic Method for Finding the Greatest or Least Value of a Function," *Computer J.*, 3, 175–184 (1960).

33. Shah, B. V., R. J. Buehler, and O. Kempthorne, "Some Algorithms for Minimizing a Function of Several Variables," *J. SIAM*, 12, No. 1, 74–92 (1964).

34. Sorenson, H. W., "Comparison of Some Conjugate Direction Procedures for Function Minimization," *J. Franklin Inst.*, 288, 421–441 (1969).

35. Wolfe, P., "The Method of Conjugate Gradients," lectures delivered at the NATO Advanced Study Institute on Mathematical Programming in Theory and Practice, Figueira da Foz, Portugal, June 1972.

第 11 章

变尺度算法

前已指出，R^n 上的可微实函数在其极小值点的邻域内可用一个二次函数来近似，基于这一假设的极小化算法，用来解无约束最优化问题通常是非常成功的. 这类算法最重要的背景之一是，用已算得的一阶导数来估计二阶导数. 在目前所知的那些估计极小化函数的 Hesse 阵或其逆阵的方法中，大多数可归入变尺度算法，此即本章所讨论的主题. 自从 1959 年 Davidon 的工作[11] 发表以来，已提出了许多其他方法. 这些方法之间的差别主要在于：从一次迭代到另一次迭代时二阶导数估计值改变的方式，以及所用线性搜索的类型和精度. 只是在最近几年，才发展了变尺度法的若干统一处理方式. 在 11.1 节中我们将通过导出整整一族算法来介绍这样一种处理方式. 这族算法中的每一种算法若是应用于具有正定阵 Q 的二次函数，则所得的搜索方向是相互共轭的，从而具有二次终结性质. 因此共轭方向将是连结前几章和本章的环节. 二次函数的另一性质被嵌入一个称为割线关系式的方程中. 它刻划了将在 11.2 节中讨论的拟 Newton 算法类的特征. 能十分成功地修改变尺度法，以适用于只有函数值可算得，甚至一阶导数也必须取近似的情况，这是 11.3 节的主题.

最后，在 11.4 节我们介绍两种其他算法，它们利用了求极小的函数的非二次性质. 在这个意义上说，它比变尺度法超出了一步. 这些方法是新的，至今还没有广泛的数值经验可供引用. 然而，它们的概念形式可能代表着发展下一代无约束极小化算法的一个方向.

这是关于无约束极小化的最后一章，无约束极小化本身就是一个重要的主题，这一章对讨论约束最优化方法的后面几章也是非常有用的.

11.1 一族变尺度算法

在 10.1 节中我们讨论了两种经典最优化方法——Newton 法和最速下降法. Newton 法的搜索方向由 $z^{k+1}=-[\nabla^2 f(x^k)]^{-1}\nabla f(x^k)$ 给定，最速下降法的搜索方向由 $z^{k+1}=-\nabla f(x^k)$ 给定. 我们已经知道，当用对称正定阵 A 定义模 x^TAx 时*，可微函数 f 在 x^k 处的方向导数取极小值的方向由 $-A^{-1}\nabla f(x^k)$ 给定. 设有椭圆形等值线的二次函数 $f(x)=\frac{1}{2}x^TQx$，其模定义为 x^TQx，通过这个例子，容易看到上述方向可能比最速下降方向更有效. 当然，这里非 Euclid 最速下降方向即为 $-Q^{-1}\nabla f(x^k)$，因为 $Q=\nabla^2 f(x^k)$，所以它也是 Newton 方向. 注意，模函数是一种特殊的尺度[3]. 因为二阶导数的计算含有特大的工作量，通常宁愿采用不用二阶导数的方法，所以就建议对计算 Newton 方向所必需的 Hesse 阵的逆阵进行近似，或等价地，可在每次迭代中改变非 Euclid 模(尺度). 在当前点 x^k，利用下式给出搜索方向 z^{k+1}：

$$z^{k+1}=-H_k^T\nabla f(x^k), \tag{11.1}$$

其中 H_k 是 $n\times n$ 阵，它从一次迭代到另一次迭代是变化的，这种极小化的迭代算法称为**变尺度法**. 注意，这个定义比在某些文献中所能找到的定义更为广泛一些. 那里限制 H_k 是对称正定的. 实际上，我们在以后将看到，这种对称正定阵 H_k 构成了变尺度算法的一个突出的子类的基础. 由于 (11.1) 与 Newton 方向的相似性，特别我们可选择矩阵 H_k，使得 R^n 上二次函数 $\frac{1}{2}x^TQx$ 的极小化在 n 次迭代后得到 $H_n^T=Q^{-1}$ 而终结；在更一般的非线性函数情形下，H_k 将趋向于最优点处的 Hesse 逆阵 $[\nabla^2 f(x^*)]^{-1}$，所以也有理由把这些方法称为拟 Newton 算法. 遗憾的是，在无约束最优化方面的研究工作者并没有使用统一的术语. 读者在文献中可发现关于"变尺度"和"拟 Newton"方法的种种不一定相同的定义. 我

---

* 译注：用 A 定义的模应为 $(X^TAX)^{1/2}$，因文中多次出现，译文未予改动.

们将限定拟 Newton 法这名字仅指变尺度法的一个子类，它的矩阵 H_k 满足所谓割线关系式，这些方法将在下节中讨论.

第一个并且也许是最著名的变尺度算法是由 Davidon[11] 提出的，随后又由 Fletcher、Powell[20] 加以简化．因而我们称它为 Davidon-Fletcher-Powell 方法或 DFP 方法．这个用来求 R^n 上可微函数 f 的极小的算法可描述如下：选取 $x^0\in R^n$ 和任一对称正定的 $n\times n$ 阵 H_0．在第 k 次迭代时，我们已有了点 x^{k-1} 和矩阵 H_{k-1}．用(11.1)计算出搜索方向 z^k．应用第8章的方法之一，寻找出 $f(x^{k-1}+\alpha_k z^k)$ 关于 α_k 的精确极小值点．令在这个极小值点处的 α_k 为 α_k^*．取 $x^k=x^{k-1}+\alpha_k^* z^k$，再用下面的修正公式计算矩阵 H_k:

$$H_k=H_{k-1}+\frac{p^k(p^k)^T}{(p^k)^T\gamma^k}-\frac{H_{k-1}\gamma^k(\gamma^k)^T H_{k-1}}{(\gamma^k)^T H_{k-1}\gamma^k}, \tag{11.2}$$

其中

$$p^k=x^k-x^{k-1},\quad \gamma^k=\nabla f(x^k)-\nabla f(x^{k-1}). \tag{11.3}$$

如果 $\|\nabla f(x^k)\|<\varepsilon$ 计算就停止；否则进行第 $k+1$ 次迭代．DFP 算法的计算经验一般是令人满意的．如果求极小的那个函数是具有正定阵 Q 的二次函数，就可建立起这个算法的一些重要性质．它具有二次终结性质，它的搜索方向关于 Q 是共轭的，在第 n 次迭代时有 $H_n=Q^{-1}$．因此这个算法既具有共轭方向法的特征，也有 Newton 型算法的特征．Sorenson[51] 已导出 DFP 方法的几个有趣的性质，并把它们同某些共轭梯度法相联系．下面我们将看到，DFP 方法只是具有同样性质的无限多种变尺度方法的一种而已.

附带提一下，这里我们可注意，在第10章中讨论过的共轭梯度算法也可看作为变尺度法，这是因为，例如由(10.96)给定的 Fletcher-Reeves 方法的搜索方向的公式可改写成

$$z^{k+1}=-H_k^T\nabla f(x^k), \tag{11.4}$$

其中

$$H_k=I-\frac{\nabla f(x^k)(z^k)^T}{(\nabla f(x^{k-1}))^T\nabla f(x^{k-1})}, \tag{11.5}$$

在下述讨论中我们将关心，从一次迭代到下次迭代时矩阵 H_k 的修改问题. 下述推导主要基于 Huang[30] 的工作，对于一大类变尺度算法，他发展了一种优美的处理方法.

因为二次终结性质似乎是希望极小化算法具有的一个性质. 并且正如我们已知道的，对于由(10.62)给定的函数，配以精确一维搜索的相互共轭的搜索方向是二次终结性质的充分条件，所以我们能要求，对于 $k=1, 2, \cdots$，所有的搜索方向应该是关于 Q 相互共轭的，即

$$(z^j)^T Q z^{k+1} = -(z^j)^T Q H_k^T \nabla f(x^k) = 0, \quad j=1, \cdots, k. \tag{11.6}$$

如果我们再假定从一点到另一点的移动如同定理 10.2 所述那样，那么

$$(z^j)^T \nabla f(x^k) = 0, \quad j=1, \cdots, k. \tag{11.7}$$

方程(11.6)和(11.7)能同时被下式给出的矩阵 H_k 满足：

$$H_k Q z^j = \omega z^j, \quad j=1, \cdots, k, \tag{11.8}$$

其中 ω 是任一实数. 假设在每次迭代中决定依照下述关系式修改矩阵 H_k：

$$H_k = H_{k-1} + \Delta H_k, \quad k=1, 2, \cdots, \tag{11.9}$$

其中矩阵 ΔH_k 称为**校正矩阵**. 我们来找出校正矩阵的一般形式. 对 $k-1$ 也写出(11.8)，再从(11.8)减去它，得

$$\Delta H_k Q z^j = 0, \quad j=1, 2, \cdots, k-1. \tag{11.10}$$

从(11.8)和(11.9)得

$$\Delta H_k Q z^k = \omega z^k - H_{k-1} Q z^k. \tag{11.11}$$

因为 $p^j = \alpha_j^* z^j$，$\gamma^j = Q p^j$，上面两个关系式成为

$$\Delta H_k \gamma^j = 0, \quad j=1, \cdots, k-1, \tag{11.12}$$

$$\Delta H_k \gamma^k = \omega p^k - H_{k-1} \gamma^k. \tag{11.13}$$

Huang 考虑的校正矩阵的一种简单形式为

$$\Delta H_k = p^k (u^k)^T + H_{k-1} \gamma^k (v^k)^T, \tag{11.14}$$

其中 u^k、v^k 是待定的 n 维向量. 首先我们看到，为了满足(11.12)和(11.13)，这些向量可这样来选择，使得

$$(u^k)^T\gamma^j=\begin{cases}0, & j=1, \cdots, k-1,\\ \omega, & j=k,\end{cases} \tag{11.15}$$

$$(v^k)^T\gamma^j=\begin{cases}0, & j=1, \cdots, k-1,\\ -1, & j=k.\end{cases} \tag{11.16}$$

在选择 u^k 和 v^k 时另一重要因素是，只利用在点 x^k 和前一点 x^{k-1} 处适用的信息. 因为我们假定搜索方向关于 Q 是共轭的，就有

$$(z^j)^TQz^k=0, \quad j=1, \cdots, k-1, \tag{11.17}$$

或者，有

$$(\gamma^j)^Tp^k=0, \quad j=1, \cdots, k-1. \tag{11.18}$$

从(11.7)推得

$$(z^j)^T\gamma^k=0, \quad j=1, \cdots, k-1, \tag{11.19}$$

再由(11.8)得

$$(z^j)^TQH_{k-1}^T\gamma^k=0, \quad j=1, \cdots, k-1, \tag{11.20}$$

或者，有

$$(\gamma^j)^TH_{k-1}^T\gamma^k=0, \quad j=1, \cdots, k-1. \tag{11.21}$$

作为(11.18)和(11.21)的结果，可以看到，如果 u^k 和 v^k 由下述形式给出，那末对于 $j=1, \cdots, k-1$, u^k 和 v^k 将满足(11.15)和(11.16)，

$$u^k=a_{11}^kp^k+a_{12}^kH_{k-1}^T\gamma^k, \tag{11.22}$$

$$v^k=a_{21}^kp^k+a_{22}^kH_{k-1}^T\gamma^k, \tag{11.23}$$

其中 a_{ii}^k 是任意参数，选得使(11.15)和(11.16)对 $j=k$ 也成立.

总之，如果矩阵 H_k 是按下列公式进行修改，

$$H_k=H_{k-1}+p^k(u^k)^T+H_{k-1}\gamma^k(v^k)^T, \tag{11.24}$$

其中 u^k 和 v^k 由(11.22)和(11.23)给定，且满足

$$(u^k)^T\gamma^k=\omega, \tag{11.25}$$

$$(v^k)^T\gamma^k=-1, \tag{11.26}$$

那么只要在每次迭代时都进行精确的线性搜索，搜索方向必将是相互共轭的. 注意修改公式也可写成

$$H_k=H_{k-1}+C_k^1A_k(C_k^2)^T, \tag{11.27}$$

其中 C_k^1 和 C_k^2 都是 $n\times 2$ 阵，

$$C_k^1=[p^k \quad H_{k-1}\gamma^k],\quad C_k^2=[p^k \quad H_{k-1}^T\gamma^k],\tag{11.28}$$

A_k 是 2×2 阵，

$$A_k=\begin{bmatrix} a_{11}^k & a_{12}^k \\ a_{21}^k & a_{22}^k \end{bmatrix}.\tag{11.29}$$

若允许数 ω 随着每次迭代而改变[48]，则讨论可进一步推广．因为到目前为止，这种推广还没有提出实际的用法，所以我们仅就 ω 是固定的情况进行讨论．采用 (11.24) 进行修改，同时满足关系式 (11.22)，(11.23)，(11.25)和(11.26)的矩阵 H_k 所组成的集合，称为 **Huang 矩阵族**.

令 $\omega=1$ 和 $a_{12}^k=a_{21}^k=0$，读者可以验证，由(11.2)给定的 DFP 方法的矩阵 H_k 就是一个 Huang 族矩阵．下述结果将表明，为了求具有正定阵 Q 的二次函数的极小，矩阵 H_k 的修改公式中参数的选择是无关紧要的，这是因为，对给定的 x^0 和 H_0，所有的 Huang 族矩阵都生成相同的点列．对于一般非线性函数，可以证明所生成的点列只依赖于 ω．在后面，我们将经常提到所用的 H_k 是 Huang 族矩阵的变尺度算法．这种算法除了以更一般的修改公式 (11.24)代替(11.2)外，其定义恰和 DFP 方法一样．下面引理属于 Shanno、Kettler[50]，但形式稍有差别.

引理 11.1

假设 f 是 R^n 上的可微实函数，又设 x^{k-1}、x^k 和 H_{k-1} 给定，满足 $(p^k)^T\nabla f(x^{k-1})\neq 0$ 和 $(p^k)^T\nabla f(x^k)=0$．那末，对所有 Huang 族矩阵 H_k，由(11.1)规定的向量 z^{k+1} 的方向都是相同的.

【证明】 根据(11.1)有

$$z^{k+1}=-H_k^T\nabla f(x^k)=-H_{k-1}^T\nabla f(x^k)-v^k(\gamma^k)^T H_{k-1}^T\nabla f(x^k)\tag{11.30}$$

$$=-H_{k-1}^T\nabla f(x^k)-(a_{21}^k p^k+a_{22}^k H_{k-1}^T\gamma^k)(\gamma^k)^T H_{k-1}^T\nabla f(x^k).\tag{11.31}$$

注意到

$$H_{k-1}^{T}\gamma^{k}=H_{k-1}^{T}\nabla f(x^{k})+\frac{p^{k}}{\alpha_{k}^{*}},\tag{11.32}$$

进行某些代数运算后，得到

$$\begin{aligned}z^{k+1}=&-\{1+a_{22}^{k}(\gamma^{k})^{T}H_{k-1}^{T}\nabla f(x^{k})\}\\&\cdot\left[I-\frac{p^{k}(\gamma^{k})^{T}}{(p^{k})^{T}\gamma^{k}}\right]H_{k-1}^{T}\nabla f(x^{k}).\end{aligned}\tag{11.33}$$

因为方括号中的表达式仅依赖于点 x^{k-1}, x^{k} 和梯度差 $\gamma^{k}=\nabla f(x^{k})-\nabla f(x^{k-1})$. 我们断定向量 z^{k+1} 的方向(但并不是它的大小)不依赖于校正矩阵的参数. ▌

引理 11.1 中的"非正交性"条件 $(p^{k})^{T}\nabla f(x^{k-1})\neq0$, 对于二次函数可以保证 $f(x^{k-1})>f(x^{k})$ (参看前一章(10.66)). 条件 $(p^{k})^{T}\nabla f(x^{k})=0$ 是沿 z^{k} 方向使 f 达到极小所必需的. 应当注意，在试图归纳地应用上述引理时，α_{k}^{*} 的选取方法必须已经确定，例如，如果 f 是具有有界水平集的严格凸函数，则 α_{k}^{*} 由下列方程的解唯一决定：

$$(z^{k})^{T}\nabla f(x^{k-1}+\alpha_{k}z^{k})=0.\tag{11.34}$$

然而，如果 f 是一般非线性函数，(11.34)的解 α_{k}^{*} 的值可能多于一个. 在这种情形下，我们可用"Curry 规则"来决定 α_{k}^{*} 的选取，即在(11.34)的解中选取能满足 $f(x^{k-1}+\alpha_{k}^{*}z^{k})<f(x^{k-1})$ 且 $|\alpha_{k}^{*}|$ 最小的那个，参看图 11.1.

现在证明，所用的 H_{k} 属于 Huang 族矩阵的任一变尺度算法，用以求具有正定阵 Q 的二次函数的极小时，所生成的点列都是相同的. 定义 n 维向量 q^{k+1} 为

$$q^{k+1}=-\left[I-\frac{p^{k}(\gamma^{k})^{T}}{(p^{k})^{T}\gamma^{k}}\right]H_{k-1}^{T}\nabla f(x^{k}).\tag{11.35}$$

定义实数 μ_{k+1} 为

$$\mu_{k+1}=\{1+a_{22}^{k}(\gamma^{k})^{T}H_{k-1}^{T}\nabla f(x^{k})\}.\tag{11.36}$$

向量 q^{k+1} 是所有 Huang 族矩阵确定的由(11.33)给出的共同方向. 因此能写成

$$z^{k+1}=\mu_{k+1}q^{k+1}.\tag{11.37}$$

现在 $p^{k+1}=\alpha_{k+1}^{*}z^{k-1}=\alpha_{k+1}^{*}\mu_{k+1}q^{k+1}$, 又由(10.64)得到

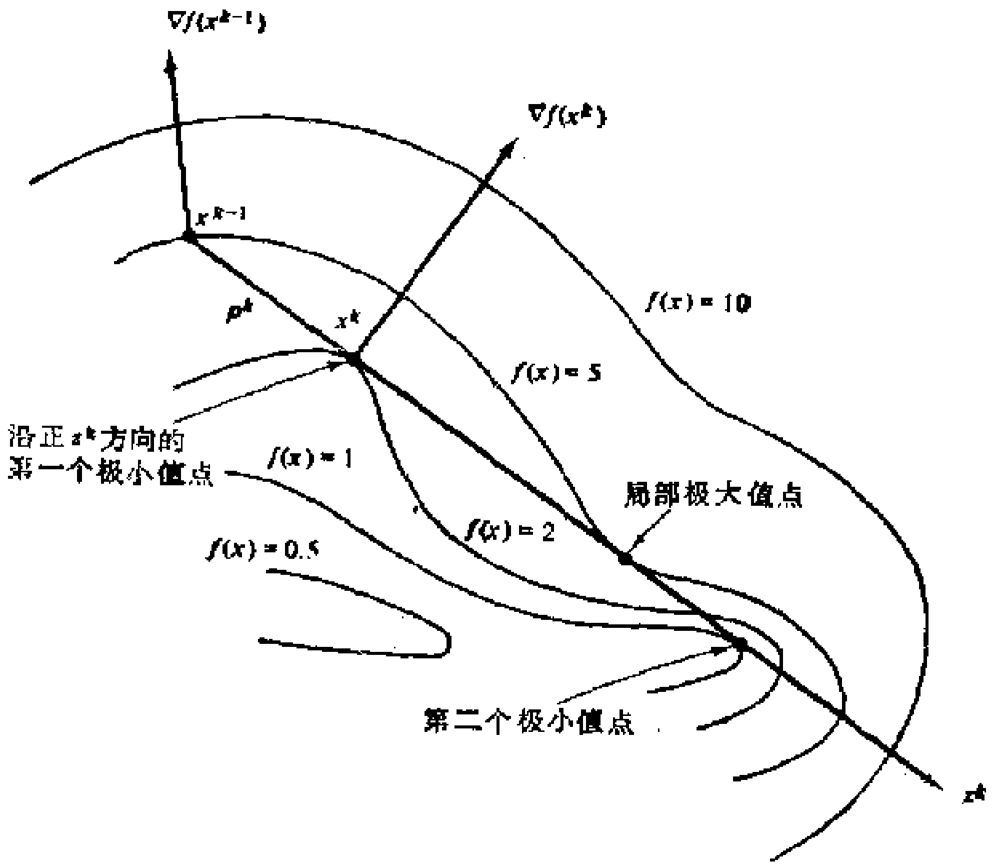


图 **11.1** 引理 11.1 的条件和 Curry 规则

$$p^{k+1}=-\frac{(q^{k+1})^T\nabla f(x^k)}{(q^{k+1})^TQq^{k+1}}\,q^{k+1}. \tag{11.38}$$

这样，下一点 x^{k+1} 不依赖于校正矩阵的参数，因而对所有的 Huang 族矩阵，x^{k+1} 都是相同的。用方向向量 q^{k+1} 可把(10.66)写成

$$f(x^k)-f(x^{k+1})=\frac{[(q^{k+1})^T\nabla f(x^k)]^2}{2(q^{k+1})^TQq^{k+1}}. \tag{11.39}$$

函数值的上述差值为正的充分必要条件是

$$(q^{k+1})^T\nabla f(x^k)\neq 0. \tag{11.40}$$

从(11.34)和(11.35)，有

$$(q^{k+1})^T\nabla f(x^k)=-(\nabla f(x^k))^TH_{k-1}\left[\nabla f(x^k)-\frac{\gamma^k(p^k)^T\nabla f(x^k)}{(p^k)^T\gamma^k}\right] \tag{11.41}$$

$$=-(\nabla f(x^k))^TH_{k-1}\nabla f(x^k), \tag{11.42}$$

应用一般的修正公式(11.22)至(11.26)，以 H_{k-2} 表示 H_{k-1}，得到

$$(\nabla f(x^k))^T H_{k-1} \nabla f(x^k) = (\nabla f(x^k))^T H_{k-2} \nabla f(x^k). \tag{11.43}$$

继续倒推，得到

$$(\nabla f(x^k))^T H_{k-1} \nabla f(x^k) = (\nabla f(x^k))^T H_0 \nabla f(x^k). \tag{11.44}$$

因此，如果(11.44)的右边非零，则(11.40)是满足的. 由此可知，对于一个下降算法，我们需要

$$(\nabla f(x^k))^T H_0 \nabla f(x^k) \neq 0, \quad k=0, \cdots, n-1. \tag{11.45}$$

并且，使任意梯度向量满足(11.45)的一个简单条件为：$\frac{1}{2}(H_0 + H_0^T)$是正定或负定的. 特别地可选取 H_0 是对称正定阵，在许多变尺度算法中就是这样选取的.

设搜索方向给定为

$$q^1 = -H_0^T \nabla f(x^0), \tag{11.46}$$

$$q^{k+1} = -\left[I - \frac{p^k(\gamma^k)^T}{(p^k)^T \gamma^k}\right] H_{k-1}^T \nabla f(x^k), \quad k=1, \cdots. \tag{11.47}$$

矩阵 H_{k-1} 可再用 H_{k-2} 来表示，继续递推，最后得到

$$q^{k+1} = -\left[I - \sum_{j=1}^{k} \frac{p^j(\gamma^j)^T}{(p^j)^T \gamma^j}\right] H_0^T \nabla f(x^k), \quad k=1, \cdots. \tag{11.48}$$

从任一点 $x^0 \in R^n$ 出发，取任意初始矩阵 H_0，使得 $(H_0 + H_0^T)/2$ 是正定或负定的，我们利用(11.46)求 f 沿着 q^1 方向的极小，找到下一点 x^1. 从 x^1 点再沿 (11.48) 给定的 q^2 方向移动. 因为 q^2 依赖于 x^0、x^1、$\nabla f(x^0)$，$\nabla f(x^1)$和 H_0，因而对于 Huang 族的所有修改公式得到的 q^2 都是相同的. 沿 q^2 方向求极小，到达 x^2. 再利用(11.48)决定一个新的方向 q^3，这样得到点列 x^0, x^1, x^2, $\cdots$, x^n. 我们正是证明了下列定理.

定理 11.2

对给定的 x^0 和 H_0，设点列 x^1, $\cdots$, x^n 是由利用 Huang 族矩阵修正公式的一个变尺度算法所产生的. 如果求极小的函数是具有对称正定阵 Q 的二次函数，那末同一点列可由任何别的 Huang 族矩阵得到.

Dixon[14,15] 把上述结果扩充到一般非线性函数，他证明了，建

立在 Huang 族矩阵基础上的变尺度算法的某种子类都产生相同的点列.

定理 11.3

假设 f 是 R^n 上可微实函数，并设 x^0 和 H_0 是给定的. 倘若对所有 k，$(p^k)^T\gamma^k\neq 0$，$a_{21}^k+a_{22}^k/\alpha_k^*\neq 0$ 且 α_k^* 为唯一选定使得 $(z^k)^T\nabla f(x^k)=0$，那末利用 Huang 族矩阵 H_k 的变尺度算法所产生的点列 x^0, x^1, …, 将只依赖于参数 ω 的选取.

【证明】 对给定的点 x^0 和矩阵 H_0，方向 z^1 对所有变尺度算法都是共同的，从而 x^1、p^1、γ^1 也如此. 根据引理 11.1，对于所有算法，z^2、x^2、p^2 和 γ^2 也是共同的. 现在假定对所有算法已存在共同点列 x^0, x^1, …, x^k. 我们将证明它们产生的 x^{k+1} 也是相同的. 根据引理 11.1，对于所有的 Huang 族矩阵，当 $k=1, 2, \cdots$ 时有

$$z^{k+1}=-\{1+a_{22}^k(\gamma^k)^TH_{k-1}^T\nabla f(x^k)\}R_kH_{k-1}^T\nabla f(x^k), \quad (11.49)$$

其中

$$R_k=I-\frac{p^k(\gamma^k)^T}{(p^k)^T\gamma^k}. \quad (11.50)$$

注意，对所有 k，有 $R_kp^k=0$. 因此，

$$\begin{aligned}R_{k+1}p^{k+1}=&-\alpha_{k+1}^*\{1+a_{22}^k(\gamma^k)^TH_{k-1}^T\nabla f(x^k)\}\\&\cdot R_{k+1}R_kH_{k-1}^T\nabla f(x^k)=0. \end{aligned}\quad (11.51)$$

因为根据(11.16)、(11.23)和假设条件，得

$$1+a_{22}^k(\gamma^k)^TH_{k-1}^T\nabla f(x^k)=-\left(a_{21}^k+\frac{a_{22}^k}{\alpha_k^*}\right)(p^k)^T\gamma^k\neq 0, \quad (11.52)$$

我们必有

$$R_{k+1}R_kH_{k-1}^T\nabla f(x^k)=0. \quad (11.53)$$

根据(11.15)、(11.22)、(11.32)和(11.50)，有

$$\begin{aligned}R_ku^k&=u^k-\frac{p^k(\gamma^k)^Tu^k}{(p^k)^T\gamma^k}=u^k-\frac{\omega p^k}{(p^k)^T\gamma^k}\\&=a_{12}^kR_kH_{k-1}^T\nabla f(x^k). \end{aligned}\quad (11.54)$$

类似地，有

$$R_k v^k = v^k - \frac{p^k(\gamma^k)^T v^k}{(p^k)^T\gamma^k} = v^k + \frac{p^k}{(p^k)^T\gamma^k}$$
$$= a_{22}^k R_k H_{k-1}^T \nabla f(x^k). \tag{11.55}$$
并且，根据(11.53)得
$$R_{k+1}u^k = \frac{\omega R_{k+1} p^k}{(p^k)^T\gamma^k}, \tag{11.56}$$
$$R_{k+1}v^k = -\frac{R_{k+1} p^k}{(p^k)^T\gamma^k}. \tag{11.57}$$
现在
$$R_k H_{k-1}^T = R_k H_{k-2}^T + R_k u^{k-1}(p^{k-1})^T + R_k v^{k-1}(\gamma^{k-1})^T H_{k-2}^T \tag{11.58}$$
$$= R_k H_{k-2}^T + \omega R_k \frac{p^{k-1}(p^{k-1})^T}{(p^{k-1})^T\gamma^{k-1}}$$
$$- R_k \frac{p^{k-1}(\gamma^{k-1})^T}{(p^{k-1})^T\gamma^{k-1}} H_{k-2}^T \tag{11.59}$$
$$= R_k R_{k-1} H_{k-2}^T + \omega R_k \frac{p^{k-1}(p^{k-1})^T}{(p^{k-1})^T\gamma^{k-1}}. \tag{11.60}$$
用递推方法，我们得到
$$R_k H_{k-1}^T = \prod_{i=1}^{k} R_i H_0^T + \omega \sum_{j=1}^{k-1}\Big(\prod_{i=j+1}^{k} R_i\Big)\frac{p^j(p^j)^T}{(p^j)^T\gamma^j}, \tag{11.61}$$
代入(11.49)，我们断定 z^{k+1} 仅依赖于参数 ω 所取的值，因此 x^{k+1} 也如此． ▌

自从 Davidon 方法提出以来，已系统地建立了许多种变尺度算法，大多数使用的校正矩阵都属于 Huang 族．我们来看一些例子．

这种校正矩阵的一个重要的子族可根据下述想法构造．在第 k 次迭代时，给定了点 x^{k-1}、x^k 和矩阵 H_{k-1}，要找下一矩阵．展开 f 的梯度
$$\nabla f(x^{k-1}) \cong \nabla f(x^k) + \nabla^2 f(x^k)(x^{k-1} - x^k) \tag{11.62}$$
或
$$\gamma^k \cong \nabla^2 f(x^k) p^k. \tag{11.63}$$
如果 $\nabla^2 f$ 是非异的，就有

$$p^k \cong [\nabla^2 f(x^k)]^{-1}\gamma^k. \tag{11.64}$$

注意，对于二次函数 f，Hesse 阵 $\nabla^2 f$ 是常数矩阵．如果它是非异的，则(11.64)对任何一步 p^k 都精确成立．因为我们感兴趣的方法并不直接利用 Hesse 逆阵，而是如同第 8 章的割线法那样仅用它的近似，所以能定义矩阵 H_k，它近似于$[\nabla^2 f(x^k)]^{-1}$，并满足所谓割线关系式

$$p^k = H_k\gamma^k. \tag{11.65}$$

基于割线关系式的各方法的若干性质将在下一节中讨论．这里仅指出，如果矩阵 H_k 是用 Huang 族校正矩阵来修正，以使(11.65)成立，则它须满足

$$p^k = H_{k-1}\gamma^k + p^k(u^k)^T\gamma^k + H_{k-1}\gamma^k(v^k)^T\gamma^k, \tag{11.66}$$

且由(11.25)和(11.26)可推断，必须选取 $\omega=1$．若进一步限制参数的选取，使(11.29)中矩阵 A_k 是对称的，即 $a_{12}^k=a_{21}^k$，则就剩下一个自由参数了，比如 a_{12}^k．

可以证明，从对称阵 H_0 出发，所有的后继矩阵 H_k 将也是对称的．它们由下式给出

$$H_k = H_{k-1} - \frac{H_{k-1}\gamma^k(\gamma^k)^T H_{k-1}}{(\gamma^k)^T H_{k-1}\gamma^k} + \frac{p^k(p^k)^T}{(p^k)^T\gamma^k} - a_{12}^k\frac{(p^k)^T\gamma^k}{(\gamma^k)^T H_{k-1}\gamma^k}\, w^k(w^k)^T, \tag{11.67}$$

其中

$$w^k = H_{k-1}\gamma^k - \frac{(\gamma^k)^T H_{k-1}\gamma^k}{(p^k)^T\gamma^k}p^k. \tag{11.68}$$

附带说明，a_{12}^k 是会选得不合适的，这时基于这些公式的算法就要失败．要求读者验证，如果

$$a_{12}^k(p^k)^T\gamma^k = \frac{(\nabla f(x^{k-1}))^T H_{k-1}\nabla f(x^{k-1})}{(\nabla f(x^k))^T H_{k-1}\nabla f(x^k)}, \tag{11.69}$$

那末 $z^{k+1}=0$.

建立在(11.67)和(11.68)给定的校正矩阵基础上的变尺度算法的一个子族是由 Broyden[7] 提出的，而后又由他继续研究[8, 9]．DFP 方法的校正矩阵，显然可由(11.67)中令 $a_{12}^k=0$ 而得到．

Broyden 的修改公式的一个重要背景是，如果 H_0 选为正定的且 $a_{12}^k \leqslant 0$，则所有 H_k 将也是正定的[7]．优先选取 H_k 是正定的，有着几个有力的理由．首先，在极小值点的邻域内，大多数函数可用一个具有正定的 Hesse 阵的二阶 Taylor 展式（二次函数）来很好地近似．因为矩阵 H_k 近似于 Hesse 逆阵．要求 H_k 是正定的，这似乎是一种好的选择．其次，本节所讨论的变尺度法在每次迭代时都要求作精确的线性搜索以找到 α_k^*．这些搜索应尽可能有效．我们已经看到，当 H_{k-1} 正定时方向 $z^k = -H_{k-1}^T \nabla f(x^{k-1})$ 是下降方向．因此在大多数情况下，只需沿着 z^k 在正的 α_k 方向上寻求 f 的极小值点．利用正定阵的另一个理由是可以避免 H_k 奇异．在第 k 次迭代中矩阵 H_k 奇异可导致大多数变尺度算法失败，这是因为如果对非零向量 $y \in R^n$，有 $H_k y = 0$，则所有后面的搜索方向 z^{k+r} 都将同 y 正交，因而它们被限制于 R^n 的一个仿射子集内．结果它们不能张成整个空间 R^n，无约束极小值点一般就不能达到．

令 $a_{12}^k = a_{21}^k$、$\omega = 1$ 和 $a_{22}^k = 0$，这就给出了满足割线关系式的一个修改公式．这时，Huang 族校正矩阵中不再有自由参数了．我们就得到(11.67)的一种特殊情形

$$H_k = H_{k-1} + \left[1 + \frac{(\gamma^k)^T H_{k-1} \gamma^k}{(p^k)^T \gamma^k}\right] \frac{p^k (p^k)^T}{(p^k)^T \gamma^k} - \frac{p^k (\gamma^k)^T H_{k-1}}{(p^k)^T \gamma^k} - \frac{H_{k-1} \gamma^k (p^k)^T}{(p^k)^T \gamma^k}. \qquad (11.70)$$

基于这个公式的算法是由 Broyden[8,9]，Fletcher[19] 和 Shanno[49] 提出的．应注意，根据定理 11.3，倘若该定理的假设条件是满足的，那末在应用于一般非线性函数时，Broyden 子族的所有算法将产生相同的点列．这个结果也已为 Dixon[14,15] 所观察到．然而上述算法中有些并不执行精确的线性搜索，因此结果可以获得 x^k 的不同序列．例如，Fletcher[19] 所提出的变尺度算法的行为和 DFP 方法并不相同．这是因为 Fletcher 方法中步长 α_k 的选法与 DFP 算法不同．在本章稍后部分，我们将回到避免线性搜索这个问题上来．

参数的另一种选择是令 $u^k=-v^k$. 由(11.25)和(11.26), 这也蕴涵着 $\omega=1$, 我们得到

$$H_k=H_{k-1}+(p^k-H_{k-1}\gamma^k)(a_{11}^k p^k+a_{12}^k H_{k-1}^T\gamma^k)^T. \tag{11.71}$$

令 $a_{11}^k=0$ 并利用(11.25), 上一公式就成为

$$H_k=H_{k-1}+\frac{(p^k-H_{k-1}\gamma^k)(\gamma^k)^T H_{k-1}}{(\gamma^k)^T H_{k-1}\gamma^k}. \tag{11.72}$$

类似地, 令 $a_{12}^k=0$, 得到

$$H_k=H_{k-1}+\frac{(p^k-H_{k-1}\gamma^k)(p^k)^T}{(p^k)^T\gamma^k}. \tag{11.73}$$

在(11.72)和(11.73)中出现的校正矩阵分别是由 Pearson 和 McCormick 提出的[40]. 令 $a_{11}^k=a_{22}^k=-a_{12}^k=-a_{21}^k$ 和 $\omega=1$, 就能得到另一校正矩阵, 结果得到下列公式:

$$H_k=H_{k-1}+\frac{(p^k-H_{k-1}\gamma^k)(p^k-H_{k-1}^T\gamma^k)^T}{(p^k-H_{k-1}^T\gamma^k)^T\gamma^k}. \tag{11.74}$$

如果选取 H_0 是对称的, 则所有的后继矩阵 H_k 将也是对称的, 并且校正矩阵的秩都为1. 基于(11.74)的算法是由 Murtagh、Sargent[36] 提出的. 他们在后来的论文[37]中也讨论了一些变尺度算法的计算背景.

例 11.1.1

我们来说明如何用变尺度算法来寻找二次函数的极小值点, 这些二次函数在以前一些例子中曾考虑过. 我们求

$$\min f(x)=\frac{3}{2}(x_1)^2+\frac{1}{2}(x_2)^2-x_1x_2-2x_1. \tag{11.75}$$

因为根据定理 **11.2**, Huang 族的每一种算法都生成同一个点列, 所以校正公式的特殊选取是无关紧要的. 不妨就取 DFP 方法. 设取初始点 $x^0=(-2, 4)^T$, H_0 是 2×2 的单位阵. 现在 $\nabla f(x^0)=(-12, 6)^T$, 这样

$$z^1=-H_0^T\nabla f(x^0)=\begin{pmatrix}12\\-6\end{pmatrix}. \tag{11.76}$$

同例 10.2.1 进行比较, 那里是用 Fletcher-Reeves 方法求同一函数的极小, 我们看到, DFP 方法同 Fletcher-Reeves 方法的第一个

方向是相同的. 对 $f(x^0+\alpha_1 z^1)$ 关于 α_1 求极小, 也同样地产生

$$x^1=\begin{pmatrix}\frac{26}{17}\\ \frac{38}{17}\end{pmatrix},\quad \nabla f(x^1)=\begin{pmatrix}\frac{6}{17}\\ \frac{12}{17}\end{pmatrix}. \tag{11.77}$$

因而

$$\begin{aligned} p^1&=\begin{pmatrix}\frac{26}{17}-(-2)\\ \frac{38}{17}-4\end{pmatrix}=\begin{pmatrix}\frac{60}{17}\\ -\frac{30}{17}\end{pmatrix},\\ \gamma^1&=\begin{pmatrix}\frac{6}{17}-(-12)\\ \frac{12}{17}-6\end{pmatrix}=\begin{pmatrix}\frac{210}{17}\\ -\frac{90}{17}\end{pmatrix}. \end{aligned} \tag{11.78}$$

下一个搜索方向由下式给出:

$$z^2=-H_1^T\nabla f(x^1), \tag{11.79}$$

其中

$$H_1=H_0+\frac{p^1(p^1)^T}{(p^1)^T\gamma^1}-\frac{H_0\gamma^1(\gamma^1)^TH_0}{(\gamma^1)^TH_0\gamma^1}. \tag{11.80}$$

于是有

$$\begin{aligned} H_1&=\begin{bmatrix}1&0\\0&1\end{bmatrix}+\frac{1}{17}\begin{bmatrix}4&-2\\-2&1\end{bmatrix}-\frac{1}{58}\begin{bmatrix}49&-21\\-21&9\end{bmatrix}\\ &=\frac{1}{986}\begin{bmatrix}385&241\\241&891\end{bmatrix}. \end{aligned} \tag{11.81}$$

这样得到

$$z^2=-\frac{1}{986}\begin{bmatrix}385&241\\241&891\end{bmatrix}\begin{pmatrix}\frac{6}{17}\\ \frac{12}{17}\end{pmatrix}=\begin{pmatrix}-\frac{9}{29}\\ -\frac{21}{29}\end{pmatrix}. \tag{11.82}$$

把 DFP 方法的 z^2 与(10.103)给定的 Fletcher-Reeves 共轭梯度算法的 z^2 比较一下, 我们断定它们的搜索向量的方向是相同的, 只是大小不同. 因而用 DFP 方法求 $f(x^1+\alpha_2 z^2)$ 关于 α_2 的极小, 所

达到的 x^2 也将是点 $(1, 1)^T$，它即为 f 的整体极小值点. 所以这两种方法产生相同点列. 这个结果并不令人意外，因为给定向量 $z^1 \in R^2$ 和矩阵 Q，共轭向量的方向是唯一确定的. 而且，我们随后还将证明对于高维的二次函数，由这两种算法产生的点列也是相同的.

读者也可验证，由 DFP 公式算得的下一个矩阵 H_2 等于 Hesse 逆阵 Q^{-1}，其中 Q 是(10.99)给定的. ▌

在上例中，矩阵 H_2 等于 Q^{-1}，这个事实可以推广. 在(11.8)中令 $k=n$，得到

$$H_nQz^j=\omega z^j, \quad j=1, \cdots, n. \tag{11.83}$$

因为搜索方向 z^j 是线性无关的，于是得到

$$H_nQ=\omega I, \tag{11.84}$$

其中 I 是 $n\times n$ 单位阵. 因此有

$$H_n=\omega Q^{-1}. \tag{11.85}$$

现在，既然在 DFP 方法中 $\omega=1$，就得到

$$H_n=Q^{-1}. \tag{11.86}$$

从共轭梯度法，诸如 Fletcher-Reeves 算法，也能得到关于 Q^{-1} 的明显表达式. 令 $z^1, \cdots, z^k$ 是 R^n 中关于 Q 相互共轭的向量. 定义

$$S_k=\sum_{j=1}^{k}\frac{z^j(z^j)^T}{(z^j)^TQz^j}. \tag{11.87}$$

于是可以证明[28]

$$S_kQz^j=\begin{cases} z^j, & j=1, \cdots, k, \\ 0, & j=k+1, \cdots, n. \end{cases} \tag{11.88}$$

因此

$$S_n=\sum_{j=1}^{n}\frac{z^j(z^j)^T}{(z^j)^TQz^j}=Q^{-1} \tag{11.89}$$

从(11.87)可得出的另一种形式为

$$S_k=\sum_{j=1}^{k}\frac{p^j(p^j)^T}{(p^j)^T\gamma^j}. \tag{11.90}$$

这样，为了计算 Q 的逆阵，可进行如下：从任一点 x^0 出发，移动到

$x^1=x^0+\alpha_1 z^1$，其中 α_1 是任一正数，并不要求它使得 $f(x^0+\alpha_1 z^1)$ 达到极小．计算 z^2，它与 z^1 关于 Q 共轭．再移动到 $x^2=x^1+\alpha_2 z^2$，其中 α_2 也是任一正数．类似地，找到 $x^3, \cdots, x^n$．计算 S_n，它就等于 Q^{-1}．

例 11.1.2

我们再次取前面例题中的二次函数．那里有共轭方向

$$z^1=\begin{pmatrix} 12 \\ -6 \end{pmatrix}, \quad z^2=\begin{pmatrix} -9 \\ -21 \end{pmatrix}. \tag{11.91}$$

从 $x^0=(-2, 4)^T$ 出发，假设我们令 $\alpha_1=1$，移动到

$$x^1=x^0+z^1=\begin{pmatrix} -2 \\ 4 \end{pmatrix}+\begin{pmatrix} 12 \\ -6 \end{pmatrix}=\begin{pmatrix} 10 \\ -2 \end{pmatrix}. \tag{11.92}$$

于是

$$p^1=\begin{pmatrix} 12 \\ -6 \end{pmatrix}, \quad \nabla f(x^1)=\begin{pmatrix} 30 \\ -12 \end{pmatrix}, \quad \gamma^1=\begin{pmatrix} 42 \\ -18 \end{pmatrix}. \tag{11.93}$$

现在令 $\alpha_2=1$，再从 x^1 移动到下一点

$$x^2=x^1+z^2=\begin{pmatrix} 10 \\ -2 \end{pmatrix}+\begin{pmatrix} -9 \\ -21 \end{pmatrix}=\begin{pmatrix} 1 \\ -23 \end{pmatrix}. \tag{11.94}$$

我们得到

$$p^2=\begin{pmatrix} -9 \\ -21 \end{pmatrix}, \quad \nabla f(x^2)=\begin{pmatrix} 24 \\ -24 \end{pmatrix}, \quad \gamma^2=\begin{pmatrix} -6 \\ -12 \end{pmatrix}. \tag{11.95}$$

因此

$$\begin{aligned} S_2&=\frac{\begin{bmatrix} 144 & -72 \\ -72 & 36 \end{bmatrix}}{(12)(42)+(6)(18)}+\frac{\begin{bmatrix} 81 & 189 \\ 189 & 441 \end{bmatrix}}{(9)(6)+(21)(12)} \\ &=\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{3}{2} \end{bmatrix}. \end{aligned} \tag{11.96}$$

S_2 便等于 Q^{-1}，Q 是由(10.99)给定的．▌

现在我们说明，应用具有标准初始方向的共轭梯度法，或应用

具有 Huang 族校正矩阵且 $H_0=I$ 的变尺度法，以求具有正定阵 Q 的二次函数 $f(x)(x\in R^n)$ 的极小时，倘若从相同的初始点 x^0 出发，那末点列 $x^1, \cdots, x^n$ 将是相同的. 对基于 Huang 族矩阵的所有变尺度算法之公共方向，重写式(11.46)和(11.48)：

$$q^1=-H_0^T\nabla f(x^0), \tag{11.97}$$

$$q^{k+1}=-\left[I-\sum_{j=1}^{k}\frac{p^j(\gamma^j)^T}{(p^j)^T\gamma^j}\right]H_0^T\nabla f(x^k), \quad k=1, 2, \cdots. \tag{11.98}$$

它们又可写成

$$q^{k+1}=-H_k^T\nabla f(x^k), \quad k=0, 1, \cdots, \tag{11.99}$$

其中

$$H_k=H_{k-1}-\frac{H_0\gamma^k(p^k)^T}{(p^k)^T\gamma^k}, \quad k=1, \cdots. \tag{11.100}$$

这个算法满足定理 10.2 的条件，因此根据(10.75)和(11.37)有

$$(q^j)^T\nabla f(x^k)=0, \quad j=1, \cdots, k. \tag{11.101}$$

由(11.98)和(11.101)，得到

$$(\nabla f(x^{j-1}))^TH_0\nabla f(x^k)-\sum_{i=1}^{j-1}\frac{(\nabla f(x^{j-1}))^TH_0\gamma^i(p^i)^T\nabla f(x^k)}{(p^i)^T\gamma^i}=0,$$
$$j=1, \cdots, k. \tag{11.102}$$

但是，对 $i=1, \cdots, j-1$ 有 $(p^i)^T\nabla f(x^k)=0$，因此

$$(\nabla f(x^{j-1}))^TH_0\nabla f(x^k)=0, \quad j=1, \cdots, k. \tag{11.103}$$

现在假设 H_0 是对称的，(11.98)可简化为

$$q^{k+1}=-\left[I-\frac{(q^k)(\nabla f(x^k))^T}{(q^k)^T\gamma^k}\right]H_0\nabla f(x^k). \tag{11.104}$$

令 $H_0=I$ 是单位阵，我们有

$$q^{k+1}=-\nabla f(x^k)+\frac{(\nabla f(x^k))^T\nabla f(x^k)}{(q^k)^T\gamma^k}q^k. \tag{11.105}$$

从(11.41)至(11.44)和(11.101)可推得

$$(q^k)^T\gamma^k=(\nabla f(x^{k-1}))^T\nabla f(x^{k-1}),$$

因此最后得到

$$q^{k+1}=-\nabla f(x^k)+\frac{(\nabla f(x^k))^T\nabla f(x^k)}{(\nabla f(x^{k-1}))^T\nabla f(x^{k-1})}q^k. \quad (11.106)$$

从而，这搜索方向与 Fletcher-Reeves 算法的情形相同.

到目前为止，仅讨论了包含"精确"线性搜索的共轭梯度法和变尺度法．即要求 α_k^* 是用求一维函数的精确极小值点来选定的那些方法．正如我们所见，这些方法都具有二次终结性质．[8]中证明了，用属于 Broyden 子族的校正矩阵且采用精确线性搜索的任何变尺度算法，去求具有正定阵 Q 的二次函数的极小时，迭代的精确次数 K 等于序列 $\nabla f(x^0)$，$QH_0\nabla f(x^0)$，$(QH_0)^2\nabla f(x^0)$，…中线性独立的 n 维向量的个数，其中 H_0 是正定的．显然，K 决不能超过 n；而依赖于初始条件，它可能小于 n．极端情况是选取 $H_0=Q^{-1}$，这蕴涵 $K=1$.

Powell[46,47] 把这个结果扩充到不是在每次迭代中都采用精确线性搜索这种有趣的情形；他证明了，如果 α_k 的选取方式使得 $f(x^{k-1}+\alpha_k z^k)<f(x^{k-1})$，那末至多进行 K 次，但不一定是相继的 K 次线性搜索后，算法就会终结，其中 K 就是上面所定义的．在每次迭代中采用精确的线性搜索这个问题是很重要的，因为这些搜索常常是变尺度算法中费时最多的部分．因此，可以避免这种精确线性搜索的变尺度法就引起极大的重视．事实上，往后我们将会看到，基于某些类型校正矩阵的若干变尺度算法不需要任何线性搜索，也显示出二次终结性质.

虽然变尺度法相当成功地被广泛应用，虽然它们也被扩充到求解约束最优化问题，但在非二次函数极小化的情形，它们的收敛性仅有少量的成果．下述诸收敛性结果属于 Powell[44]．证明十分长，因而略去了.

定理 11.4

设 f 是 R^n 上的二次连续可微实函数，对于某个 $m>0$ 和 R^n 中每个 x 与 y，满足

$$y^T\nabla^2 f(x)y\geqslant m\|y\|^2, \quad (11.107)$$

那末对任何 $x^0\in R^n$，满足 $f(x)\leqslant f(x^0)$的 x 点的集合必是紧致的，

且由 DFP 算法产生的序列 $x^0, x^1, x^2, \cdots$ 收敛于 f 的唯一极小值点 x^*.

在加于 f 的条件下，从定理 11.3 和定理 11.4 可推得，用 Huang 族校正矩阵且 $\omega=1$ 的任何变尺度算法，将收敛于 f 的极小值点. 注意，上一定理中关于 f 的条件蕴涵着 f 是一个严格凸函数，且在它的每一点 $x\in R^n$ 上的 Hesse 阵的特征值都有正下界 m. 下一定理论及 DFP 算法的收敛速率.

定理 11.5

设定理 11.4 的假设条件成立，且对所有满足 $f(x)\leqslant f(x^0)$ 的 x 存在一个常数 L，使得

$$\left|\frac{\partial^2 f(x)}{\partial x_i \partial x_j}-\frac{\partial^2 f(x^*)}{\partial x_i \partial x_j}\right|\leqslant L\|x-x^*\|,$$

$$i=1, \cdots, n, \quad j=1, \cdots, n. \tag{11.108}$$

那末由 DFP 算法产生的点列 $x^0, x^1, \cdots$ 在(8.21)的意义下超线性收敛；即

$$\lim_{k\to\infty}\frac{\|x^{k+1}-x^*\|}{\|x^k-x^*\|}=0. \tag{11.109}$$

由此还可得出，每一种具有 $\omega=1$ 的变尺度算法也是超线性收敛的. 我们强调这一点，因为 Powell[45] 曾猜测，前面定理的冗长的证明，可能由于使用 $\omega=1$ 的另一不同的 Huang 族校正矩阵来代替 DFP 公式而缩短. 另一方面，上述定理的证明，对建立 $\omega\neq1$ 的其他变尺度法的收敛性可能是有用的.

Powell[46] 已把定理 11.4 的结果扩充到具有有界 Hesse 阵的实值凸函数上.

定理 11.6

设 f 是 R^n 上的二次连续可微的凸函数，对所有 x 满足

$$\|\nabla^2 f(x)\|\leqslant M. \tag{11.110}$$

又设满足 $f(x)\leqslant f(x^0)$ 的 x 的集合是紧致的. 那末 DFP 方法产生的点列 $x^0, x^1, \cdots$ 的每一凝聚点都是 f 的极小值点.

Goldfarb[23] 也讨论了变尺度法的收敛性. 他证明了，每次迭

代都采用精确线性搜索的每种变尺度算法，对于满足式(10.120)，即 $\nabla^2 f(x)$ 的特征值是有界的严格凸函数 f，如果矩阵 H_k 都是正定的且它们的特征值都介于两个正数之间，那末算法将收敛于 f 的唯一极小值点．他也找到了关于某一类变尺度法使矩阵 H_k 有这种性质的充分条件．

11.2 拟 Newton 法

我们刚才基于共轭方向和精确线性搜索，导出了具有二次终结性质的一族变尺度算法．这里我们将讨论另外几种变尺度算法，它们的成功应用，并不需要精确的线性搜索．本节的许多具体公式既可以根据上一节论据推导出来，也可以根据本节所用的论据推导出来，从这个意义说，本节的矩阵修改公式与上一节的有关公式有大量是重复的．虽然前节的变尺度法特别适合于求函数的极小值，但本节的一些方法原来却是为解非线性方程组而导出的，这方程组并不一定是令函数的梯度为零而得出的．既然我们的主要目标是最优化，我们将继续使用和前面相同的概念．对非线性方程组感兴趣的读者，可容易地改写下列全部结果．

这里所讨论的方法对第 10 章所述的 Newton 法提出某种修正．仍设我们希望求 R^n 上可微函数 f 的极小．为求解由 $\nabla f(x)=0$ 给定的 n 个变量和 n 个方程的方程组，Newton 法是由下列迭代公式所支配：

$$x^{k+1}=x^k-[\nabla^2 f(x^k)]^{-1}\nabla f(x^k),\quad k=0,\ 1,\ \cdots. \tag{11.111}$$

这方法的两个主要缺点是，(a) 为求 Hesse 逆阵需要大量计算，(b) 方法不稳定，可能收敛到局部极大值点，或者当给定的初始值估计很坏时可能不收敛[6]．下面所述的方法在每次迭代中求 Hesse 逆阵的近似，而不需要求 f 的二阶偏导数的值．假设有一非异的 $n\times n$ 阵 B_k，它近似于 $\nabla^2 f(x^k)$．那末可选取由下式给定的方向 z^{k+1}：

$$z^{k+1}=-B_k^{-1}\nabla f(x^k), \tag{11.112}$$

沿此方向移动到下一点 x^{k+1}：

$$x^{k+1}=x^k+\alpha_{k+1}z^{k+1}, \tag{11.113}$$

其中 α_{k+1} 是一实数，以这样的方式选取，使得迭代最后能收敛于所需要的解．一般地可定义

$$x=x^k+\theta z^{k+1}, \tag{11.114}$$

并把 $\nabla f(x)=\nabla f(x^k+\theta z^{k+1})$ 看作单变量 θ 的函数．现在

$$\frac{d(\partial f(x)/\partial x_j)}{d\theta}=\sum_{l=1}^{n}\frac{\partial^2 f(x)}{\partial x_j\partial x_l}\frac{dx_l}{d\theta}=\sum_{l=1}^{n}\frac{\partial^2 f(x)}{\partial x_j\partial x_l}z_l^{k+1},$$

$$j=1,\ \cdots,\ n. \tag{11.115}$$

这样，如果寻找在 x^{k+1} 处 Hesse 阵的近似值，上面这个关系式表明它能借助于 $d(\partial f(x^{k+1})/\partial x_j)/d\theta$ 的近似值得到．利用 (11.115)，这样一个近似可写成

$$\frac{\partial f(x^k+\theta z^{k+1})}{\partial x_j}\cong\frac{\partial f(x^{k+1})}{\partial x_j}+\sum_{l=1}^{n}\frac{\partial^2 f(x^{k+1})}{\partial x_j\partial x_l}z_l^{k+1}(\theta-\alpha_{k+1}),$$

$$j=1,\ \cdots,\ n. \tag{11.116}$$

如果已求得了 ∇f 在 $x^k+\theta z^{k+1}$ 和 x^{k+1} 这两点处之值，则由(11.116)可写出

$$\frac{\partial f(x^{k+1})}{\partial x_j}-\frac{\partial f(x^k+\theta z^{k+1})}{\partial x_j}\cong\sum_{l=1}^{n}\frac{\partial^2 f(x^{k+1})}{\partial x_j\partial x_l}z_l^{k+1}(\alpha_{k+1}-\theta),$$

$$j=1,\ \cdots,\ n, \tag{11.117}$$

或

$$\nabla f(x^{k+1})-\nabla f(x^k+\theta z^{k+1})\cong\nabla^2 f(x^{k+1})(\alpha_{k+1}-\theta)z^{k+1}. \tag{11.118}$$

在本节方法背后的基本思想是，选择 f 的 Hesse 阵的下一次近似值 B_{k+1}，使之满足方程

$$\nabla f(x^{k+1})-\nabla f(x^k+\theta z^{k+1})=B_{k+1}(\alpha_{k+1}-\theta)z^{k+1}. \tag{11.119}$$

通常选取 θ 等于零，于是得到

$$\nabla f(x^{k+1})-\nabla f(x^k)=\gamma^{k+1}=B_{k+1}(x^{k+1}-x^k)=B_{k+1}p^{k+1}. \tag{11.120}$$

既然对于迭代过程来说，比起 B_{k+1} 本身，更需要的是 B_{k+1}^{-1}，我们定义 $H_k=B_k^{-1}$，(11.120)就变成了割线关系式(11.65)

$$p^{k+1}=H_{k+1}\gamma^{k+1}. \tag{11.121}$$

采用割线关系式来修改矩阵 H_k 的变尺度法称为拟**Newton**法[6]。在这些方法中，有

$$z^{k+1} = -H_k \nabla f(x^k), \tag{11.122}$$

其中 H_k 的转置号去掉了，x^{k+1} 由(11.113)给定，而 α_{k+1} 之值并不一定要使 $f(x^k+\alpha z^{k+1})$ 达到极小.

现在转到如何修改 H_k 这个问题. 如同上节一样，假定

$$H_k = H_{k-1} + \Delta H_k. \tag{11.123}$$

于是由割线关系式得

$$\Delta H_k \gamma^k = p^k - H_{k-1}\gamma^k. \tag{11.124}$$

如果令

$$\Delta H_k = \frac{(p^k - H_{k-1}\gamma^k)(y^k)^T}{(y^k)^T \gamma^k}, \tag{11.125}$$

其中 y^k 是使得 $(y^k)^T\gamma^k \neq 0$ 的任意 n 维向量，那末 (11.124) 式成立，从而 (11.121) 也成立. 此外，(11.125) 右边的矩阵在使(11.124)能满足的所有矩阵中，其秩是最小的. 当然，(11.125)中矩阵 ΔH_k 的秩为 1. 因此一般的秩 1 修改公式可写成[7, 10]

$$H_k = H_{k-1} + \frac{(p^k - H_{k-1}\gamma^k)(y^k)^T}{(y^k)^T \gamma^k}, \tag{11.126}$$

其中 y^k 受限制于$(y^k)^T\gamma^k \neq 0$.

现在考虑几种更特殊的拟 Newton 法，它们所用的修正矩阵是上述一般秩 1 矩阵的特例. 在 Wolfe[54] 和 Barnes[2] 的割线法中，向量 y^k 的选法如下：对于 $k<n$，使得 y^k 满足

$$(y^k)^T\gamma^i = 0, \quad i=1, \cdots, k-1. \tag{11.127}$$

因为这个关系式还不能唯一确定 y^k，所以建议 y^k 取为 $\gamma^1, \cdots, \gamma^k$ 的线性组合，且正交于 $\gamma^1, \cdots, \gamma^{k-1}$. 这一步骤可采用 Gram-Schmidt 正交化方法来完成. 如果 $k \geqslant n$，则选取 y^k 使得

$$(y^k)^T\gamma^i = 0, \quad i=k-n+1, \cdots, k-1. \tag{11.128}$$

容易证明，如果 y^k 满足上述关系式，则对于 $k \leqslant n$，

$$p^i = H_k\gamma^i, \quad i=1, \cdots, k. \tag{11.129}$$

在割线法的原来形式中，它实际上是对 Hesse 阵而不是对 Hesse

逆阵作近似. 现在的形式更为有效,因为它不必去求逆阵.

Murtagh、Sargent[36] 证明了下述结果, 这些结果蕴涵割线法具有二次终结性质.

定理 11.7

给定使(11.129)成立的各步 p^i, $i=1, \cdots, k$, 且 H_k 非异, 则对于二次函数 f, 由

$$p^{k+1}=-H_k\nabla f(x^k) \tag{11.130}$$

给定的这一步,或者是线性独立于以前各步,或者它已到达 f 的逗留点.

【证明】 设 p^{k+1} 线性相关于 $p^1, \cdots, p^k$, 则存在不全为零的数 $\beta_1, \cdots, \beta_k$, 使得

$$p^{k+1}=\sum_{i=1}^{k}\beta_i p^i. \tag{11.131}$$

因为 f 是二次函数, 我们有 $f(x)=a+b^Tx+\frac{1}{2}x^TQx$, 从而

$$\gamma^i=Qp^i. \tag{11.132}$$

因此

$$\gamma^{k+1}=Qp^{k+1}=\sum_{i=1}^{k}\beta_i Qp^i=\sum_{i=1}^{k}\beta_i\gamma^i. \tag{11.133}$$

从(11.129)、(11.131)和(11.133), 可知

$$p^{k+1}=\sum_{i=1}^{k}\beta_i H_k\gamma^i=H_k\sum_{i=1}^{k}\beta_i\gamma^i=H_k\gamma^{k+1}. \tag{11.134}$$

再由(11.130)和(11.134), H_k 又是非异的,就得到

$$-\nabla f(x^k)=\gamma^{k+1}=\nabla f(x^{k+1})-\nabla f(x^k), \tag{11.135}$$

也就是

$$\nabla f(x^{k+1})=0. \tag{11.136}$$

】

推论 11.6

如果 H_k 非异且满足(11.129), 则二次函数的逗留点至多用 $n+1$ 步就可达到.

【证明】 因为至多只能有 n 个线性独立的步, 所以由前一定理直接就可推出. 】

割线法的二次终结性质，也可通过说明割线法所产生的矩阵 $H_0, H_1, \cdots, H_n$ 逐渐接近于 Q^{-1} 使 $H_n=Q^{-1}$，而获得证明。这样，在 n 步之后，下一步就是原来的 Newton 步，一定给出解。注意，在二次函数的情形，步长大小对于有限终结性来说是无关紧要的。

对一般非线性函数，割线法可能需要多于 n 步才能满足某种收敛准则。这时，y^k 由(11.128)选取，经验表明相继的搜索方向会成为线性相关，因而必须对(11.128)作一些修改。

在基于(11.120)的 Broyden 1965 方法[6] 中，采用了稍有不同的途径。他注意到，(11.120)把 $\nabla f(x^k)$ 的改变与位置 x^k 在 p^{k+1} 方向上的改变联系起来了，但关于 $\nabla f(x^k)$ 在不同于 p^{k+1} 的方向上的改变没有提供信息。Broyden 提议，设 γ^{k+1} 是正交于 p^{k+1} 的一个方向，B_{k+1} 的选取应使由 B_{k+1} 所预估的 $\nabla f(x^k)$ 在 γ^{k+1} 上的改变，与由 B_k 所给定的改变相同。形式地，有

$$B_{k+1}\gamma^{k+1}=B_k\gamma^{k+1},\quad (p^{k+1})^T\gamma^{k+1}=0. \tag{11.137}$$

读者容易验证，由

$$B_k=B_{k-1}+\frac{(\gamma^k-B_{k-1}p^k)(p^k)^T}{(p^k)^Tp^k} \tag{11.138}$$

给定的 B_k 同时满足(11.120)和(11.137)。因为我们需要的是 $B_k^{-1}=H_k$ 而不是 B_k 本身，所以对非异的 $n\times n$ 阵 A 和 n 维向量 u 和 v，可以利用 Sherman-Morrison 公式[29]：

$$(A+uv^T)^{-1}=A^{-1}-\frac{A^{-1}uv^TA^{-1}}{1+v^TA^{-1}u}. \tag{11.139}$$

令 $A=B_{k-1}$，$A^{-1}=H_{k-1}$，$u=(\gamma^k-B_{k-1}p^k)/(p^k)^Tp^k$，$v=p^k$，由上两个方程得到

$$H_k=H_{k-1}-\frac{(H_{k-1}\gamma^k-p^k)(p^k)^TH_{k-1}}{(p^k)^TH_{k-1}\gamma^k}. \tag{11.140}$$

这样，Broyden 方法利用由(11.122)给出的搜索方向 z^{k+1}，而步长 $p^{k+1}=x^{k+1}-x^k=\alpha_{k+1}z^{k+1}$，其中 α_{k+1} 通常选得使

$$\|\nabla f(x^{k+1})\|\leqslant\|\nabla f(x^k)\|. \tag{11.141}$$

这里并不存在特别的理由，应用精确的线性搜索来确定步长，因而 α_{k+1} 也可取为1. 这个方法并不具有二次终结性质.

Broyden[10] 注意到，如果 B_k 是 $\nabla^2 f(x^k)$ 的好的近似，则配以 $\alpha_k=1$ 的算法通常是稳定的. 然而如果 B_k 不是 f 的 Hesse 阵的好的近似，这方法就成为不稳定的. 这算法的收敛性被认为优于割线法，也由 Broyden[10] 作了研究，他证明了，在适当的假定下用于求二次函数的极小时，他的方法的收敛速率是超线性的. 这些结果后来由 Dennis[13] 扩充到一般的非线性函数.

例 11.2.1

我们对已熟悉的二次函数

$$f(x)=\frac{3}{2}(x_1)^2+\frac{1}{2}(x_2)^2-x_1x_2-2x_1, \tag{11.142}$$

进行少量迭代来说明 Broyden 1965 方法.

假设 $H_0=I$，并对所有 k 取 $\alpha_k=1$. 从 $x^0=(-2,\ 4)^T$ 这点出发搜索，在 x^0 处 $f(x^0)=26$，$\nabla f(x^0)=(-12,\ 6)^T$. 下一点为

$$x^1=x^0-H_0\nabla f(x^0)=\begin{pmatrix}10\\-2\end{pmatrix}. \tag{11.143}$$

在这点处 $f(x^1)=152$，$\nabla f(x^1)=(30,\ -12)^T$. 由此得

$$\gamma^1=\begin{pmatrix}42\\-18\end{pmatrix},\quad p^1=\begin{pmatrix}12\\-6\end{pmatrix}, \tag{11.144}$$

$$H_1=\begin{bmatrix}1&0\\0&1\end{bmatrix}-\frac{\left[\begin{pmatrix}42\\-18\end{pmatrix}-\begin{pmatrix}12\\-6\end{pmatrix}\right](12,\ -6)\begin{bmatrix}1&0\\0&1\end{bmatrix}}{(12,\ -6)\begin{bmatrix}1&0\\0&1\end{bmatrix}\begin{pmatrix}42\\-18\end{pmatrix}}, \tag{11.145}$$

即

$$H_1=\begin{bmatrix}0.41176&0.29412\\0.23529&0.88235\end{bmatrix}. \tag{11.146}$$

下一步为

$$p^2=-H_1\nabla f(x^1)=-\begin{bmatrix}0.41176 & 0.29412\\ 0.23529 & 0.88235\end{bmatrix}\begin{pmatrix}30\\ -12\end{pmatrix}$$

$$=\begin{pmatrix}-8.82353\\ 3.52941\end{pmatrix}. \tag{11.147}$$

因此

$$x^2=\begin{pmatrix}10\\ -2\end{pmatrix}+\begin{pmatrix}-8.82353\\ 3.52941\end{pmatrix}=\begin{pmatrix}1.17647\\ 1.52941\end{pmatrix}. \tag{11.148}$$

在这点处 $f(x^2)=-0.90657$ 和 $\nabla f(x^2)=(0,0.35294)^T$. 这样,

$$\gamma^2=\begin{pmatrix}0\\ 0.3529\end{pmatrix}-\begin{pmatrix}30\\ -12\end{pmatrix}=\begin{pmatrix}-30\\ 12.3529\end{pmatrix}, \tag{11.149}$$

$$H_2=H_1-\frac{(H_1\gamma^2-p^2)(p^2)^TH_1}{(p^2)^TH_1\gamma^2} \tag{11.150}$$

$$=\begin{bmatrix}0.41498 & 0.29352\\ 0.24494 & 0.88058\end{bmatrix}. \tag{11.151}$$

由此得

$$p^3=-H_2\nabla f(x^2)=\begin{pmatrix}-0.10359\\ -0.31079\end{pmatrix}, \tag{11.152}$$

$$x^3=\begin{pmatrix}1.17647\\ 1.52941\end{pmatrix}-\begin{pmatrix}0.10359\\ 0.31079\end{pmatrix}=\begin{pmatrix}1.07288\\ 1.21862\end{pmatrix}. \tag{11.153}$$

在这点处,有

$$f(x^3)=-0.98407, \qquad \nabla f(x^3)=\begin{pmatrix}0.00002\\ 0.14574\end{pmatrix}. \tag{11.154}$$

可以看出, 第一步采取的是最速下降方向(因为 H_0 取为 I). 这一步之后的收敛情况是好的, 但不存在二次终结性质. 如果计算 H_3, 得到

$$H_3=\begin{bmatrix}0.49560 & 0.49981\\ 0.48683 & 1.49945\end{bmatrix}, \tag{11.155}$$

这个矩阵是 Q^{-1} 的一个好的近似, Q^{-1} 是

$$Q^{-1}=\begin{bmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{3}{2}\end{bmatrix}. \qquad (11.156)$$

注意，虽然选取 H_0 是对称的，但后继的矩阵可以是不对称的，如在 H_1 和 H_2 中明显看到的那样．当然，当 H_k 收敛于 Q^{-1} 时，它们必将再趋向于对称． ▌

这个例子表明，在 Broyden 1965 方法中（在割线法中也同样），矩阵 H_k 并不一定是对称的，甚至在 H_0 为对称时 H_k 亦未必为对称．既然 H_k 被要求为 Hesse 逆阵的近似，而 Hesse 逆阵是对称阵，因此，寻找一个使得所有 H_k 为对称且满足割线关系式的修改公式，是合情合理的．

这种修改公式是Broyden[7]，Davidon[12]，Murtagh、Sargent[36]和 Powell[42] 所提出的**对称秩 1** 算法的基础．在(11.126)中，取 y^k 为

$$y^k=p^k-H_{k-1}\gamma^k, \qquad (11.157)$$

则矩阵 H_k 为

$$H_k=H_{k-1}+\frac{(p^k-H_{k-1}\gamma^k)(p^k-H_{k-1}\gamma^k)^T}{(p^k-H_{k-1}\gamma^k)^T\gamma^k}. \qquad (11.158)$$

显然，对于任何对称的 H_{k-1}，H_k 也将是对称的．此外，建立在这个矩阵修改公式基础上的算法具有二次终结性质，而不必进行精确的线性搜索．我们用下列事实证明这结果：在具有非异阵 Q 的二次函数的情形，只要内积 $(p^k-H_{k-1}\gamma^k)^T\gamma^k$ 无一为零，则关系式(11.129)成立，即

$$p^i=H_k\gamma^i, \quad i=1, \cdots, k. \qquad (11.159)$$

假如

$$p^i=H_{k-1}\gamma^i, \quad i=1, \cdots, k-1, \qquad (11.160)$$

成立．现在，对于 $i=1, \cdots, k-1$，有

$$(p^k-H_{k-1}\gamma^k)^T\gamma^i=(p^k)^T\gamma^i-(H_{k-1}\gamma^k)^T\gamma^i \qquad (11.161)$$

$$=(p^k)^T\gamma^i-(\gamma^k)^Tp^i \qquad (11.162)$$

$$=(\gamma^k)^TQ^{-1}\gamma^i-(\gamma^k)^TQ^{-1}\gamma^i=0, \qquad (11.163)$$

其中 $p^i=Q^{-1}\gamma^i$ 对 $i=1, \cdots, k$ 成立. 因此, 由(11.158)、(11.160)至(11.163), 得到

$$H_k\gamma^i=H_{k-1}\gamma^i=p^i, \quad i=1, \cdots, k-1. \qquad (11.164)$$

又因为 H_k 满足割线关系式, 就可断定(11.129)成立. 再借助于推论 11.6, 就得到二次终结性质. 注意, 在证明二次终结性质时, 沿搜索方向 z^k 的一步 p^k 所赖以决定的 α_k 的取法是无关紧要的. 要求读者证明下列有趣的结果: 如果 α_k 的选取使得具有正定阵 Q 的二次函数达到极小, 那末对称秩 1 算法的搜索方向 $z^1, \cdots, z^n$ 关于 Q 确实是共轭的.

既然修改公式(11.74)和(11.158)是一致的, 所以用(11.158)修改的对称秩 1 方法的矩阵 H_k 也是一个 Huang 族矩阵. 取 $\omega=1$, 从定理 11.3 可知, 在求 R^n 上可微函数的极小时, 配以精确线性搜索的对称秩 1 算法所产生的点列, 将和把 Huang 族中 A_k 限定为对称阵且 H_k 满足割线关系式的那种 Broyden 方法所产生的点列相同. 基于一般 Huang 族矩阵的算法仅当每次迭代都采用精确线性搜索时才呈现出二次终结性质. 而对称秩 1 方法无须线性搜索在二次函数上也能终结. 虽然用对称秩 1 方法求二次函数的极小时, 所需计算量小于其他采用线性搜索的变尺度算法, 但是对于非二次函数是否仍具有这种优点还不清楚.

Broyden[10] 列出选取步长 α_k 的三种常用办法, 并考察了它们的效果:

1. 精确的线性搜索, 就是使得 $(z^k)^T\nabla f(x^k)=0$.
2. 步长的选取使得 $\|\nabla f\|$ 沿 z^k 方向达到极小或有所减小.
3. "Newton"步, 即对所有 k, $\alpha_k=1$.

当然也可设计步长的其他选取办法, 诸如在第 10 章中提到的保证"函数值充分减小"的办法, 或者在 Goldstein、Price[24] 方法中的有关选取办法. Dixon[16] 进行了各种选取办法的数值比较, 在另外的选取办法中, 他检验了每次迭代中用单个二次逼近或三次逼近来代替完全的一维极小化的效果.

既然拟 Newton 法的收敛性是通过对 Hesse 逆阵逐步逼近而

获得的，Broyden[8,9,10] 提出用考察某种矩阵序列的性态来分析拟 Newton 法的收敛性，这序列中每个矩阵指出了矩阵 H_k 同 Hesse 逆阵的接近程度. 假设我们求 R^n 上具有非异阵 Q 的二次函数的极小，定义

$$E_k = Q^{-1}B_k - I, \tag{11.165}$$

$$G_k = H_k Q - I. \tag{11.166}$$

如同前面一样，其中 $B_k = H_k^{-1}$. 显然，E_k 和 G_k 成为零矩阵的充分必要条件为 $H_k = Q^{-1}$. 我们可把矩阵 E_k 或 G_k 的秩或它们的某种模逐步减小到零，来接近 $H_k = Q^{-1}$ 这情形. [10]中说明了割线法和对称秩 1 算法都是"减秩的"而 Broyden 1965 方法则是"减模的".

到目前为止，我们主要涉及秩 1 的拟 Newton 算法. 由 (11.67) 和 (11.68) 给定的 Broyden 1967 年的修改关系式是作为一种特殊的 Huang 族矩阵导出的. 但是原来的叙述是循着本节所用的拟 Newton 方式，并需要一个秩 2 校正矩阵 ΔH_k. 为了方便起见，重写 Broyden 1967 修改关系式，它有一个自由参数，

$$H_k = H_{k-1} - \frac{H_{k-1}\gamma^k(\gamma^k)^T H_{k-1}}{(\gamma^k)^T H_{k-1}\gamma^k} + \frac{p^k(p^k)^T}{(p^k)^T\gamma^k} - a_{12}^k \frac{(p^k)^T\gamma^k}{(\gamma^k)^T H_{k-1}\gamma^k} w^k(w^k)^T, \tag{11.167}$$

其中

$$w^k = H_{k-1}\gamma^k - \frac{(\gamma^k)^T H_{k-1}\gamma^k}{(p^k)^T\gamma^k} p^k. \tag{11.168}$$

令

$$a_{12}^k = -\frac{1}{(p^k)^T\gamma^k}, \tag{11.169}$$

代入(11.167)，可得到 Broyden[9] 提出的一个特殊公式，它就是上节讨论过的式(11.70)

$$H_k = H_{k-1} + \left[1 + \frac{(\gamma^k)^T H_{k-1}\gamma^k}{(p^k)^T\gamma^k}\right]\frac{p^k(p^k)^T}{(p^k)^T\gamma^k} - \frac{p^k(\gamma^k)^T H_{k-1}}{(p^k)^T\gamma^k} - \frac{H_{k-1}\gamma^k(p^k)^T}{(p^k)^T\gamma^k}. \tag{11.170}$$

正如已注意到的，当配以精确线性搜索且 H_0 取为正定阵而用于求具有正定阵 Q 的二次函数的极小时，a_{12}^k 的上述取法保证了上式给出的所有 H_k 将也是正定的. 这个修改公式在文献中通常称为 Broyden-Fletcher-Shanno 公式或 BFS 公式. Fletcher[19] 证明了，在 DFP 和 BFS 修改公式之间的一个有趣的关系. 设 H_k 是用 DFP 公式修改得出的. 令 $B_{k-1}^{-1}=H_{k-1}$, $B_k^{-1}=H_k$, 我们得到

$$B_k=B_{k-1}+\left[1+\frac{(p^k)^T B_{k-1}p^k}{(\gamma^k)^T p^k}\right]\frac{\gamma^k(\gamma^k)^T}{(\gamma^k)^T p^k} -\frac{\gamma^k(p^k)^T B_{k-1}}{(\gamma^k)^T p^k}-\frac{B_{k-1}p^k(\gamma^k)^T}{(\gamma^k)^T p^k}. \tag{11.171}$$

即 B_k 是按照和 BFS 公式类似的公式来修改的，其中用 B_{k-1} 代替 H_{k-1}，且 p^k 与 γ^k 互换. 相反，如果 H_k 是用 BFS 公式修改的，则 B_k 相应地是按照和 DFP 公式类似的公式修改的. 由于这些关系，BFS 公式有时也称为相补 DFP 公式. 在 Murray [35] 中可找到从各种不同的拟 Newton 法导出的 Hesse 阵的各种近似阵 B_k 之间的一般关系.

DFP 公式和 BFS 公式的另一个共同性质是，二者都包含有建立在它们基础上的、放弃精确线性搜索的方法，而在每次迭代中使用某种减小函数值的步骤，以保持矩阵 H_k 的正定性. 这种方法以上面讲到过的"减模"方式收敛到二次函数的极小值点.

Fletcher[19] 证明了，相应于这两种方法的变尺度矩阵 H_k^{DFP} 和 H_k^{BFS} 的任何凸组合，也具有上述性质，这时的矩阵 H_k 为

$$H_k=q_1H_k^{DFP}+q_2H_k^{BFS}, \tag{11.172}$$

其中 $q_1\geqslant 0$、$q_2\geqslant 0$、$q_1+q_2=1$. 实际上，Fletcher 推荐了一种算法，它的变尺度用 DFP 公式或 BFS 公式来修正，取决于下述检验. 如果

$$(p^k)^T\gamma^k\geqslant(\gamma^k)^T H_{k-1}\gamma^k, \tag{11.173}$$

则采用 BFS 公式修正；否则就用 DFP 公式. 步长的决定类似于上一章 Goldstein-Price 算法中的方式. 取一小的正数 $\delta<\frac{1}{2}$，在每次迭代中，计算

$$g(x^{k-1},\ \alpha_k)=\frac{f(x^{k-1})-f(x^{k-1}+\alpha_k z^k)}{\alpha_k(z^k)^T\nabla f(x^{k-1})}\tag{11.174}$$

在 $\alpha_k=1$ 时的值. 如果 $g(x^{k-1},\ 1)\geqslant\delta$, 则取 $x^k=x^{k-1}+z^k$; 否则减小 α_k 直至满足 $g(x^{k-1},\ \hat{\alpha}_k)\geqslant\delta$, 并取 $x^k=x^{k-1}+\hat{\alpha}_k z^k$.

一些研究者认为 Hesse 逆阵的近似 H_k 的正定性是一个成功的变尺度算法的基本性质之一. 如果 H_k 在每次迭代中都保持正定, 那末对于求极小的函数来说, 函数值的减小就有了保证. 然而, 甚至理论上能保持矩阵 H_k 正定性的算法, 如 DFP 方法, 也会出现数值方面的困难, 以致使这些矩阵变成奇异或不定[1]. 前已指出, 可以证明, 一旦计算到一个奇异阵 H_k, 后继的搜索方向 z^{k+1}, z^{k+2}, …将不能张成定义求极小的那个函数的整个空间. Gill、Murray[21] 提出了对称秩 2、拟 Newton 算法的改进形式, 使诸如 DFP 或 BFS 方法, 它不但在理论上保证正定性, 而且在实际上也能保持这性质. 他们用解下列线性方程组来求得搜索方向 z^k:

$$B_k z^{k+1}=-\nabla f(x^k),\tag{11.175}$$

其中 B_k 是 f 的 Hesse 阵的近似. 如果 B_k 是对称正定阵, 它可写成

$$B_k=L_kD_kL_k^T,\tag{11.176}$$

其中 $L_k=[l_{ij}^k]$是对角元素均为 1 的下三角阵, $D_k=[d_{ij}^k]$ 是对角元素均为正数的对角阵. 式(11.176) 的右边称为 B_k 的 Cholesky 分解[53]. 把(11.176)代入(11.175)中, 得到

$$L_kD_kL_k^Tz^{k+1}=-\nabla f(x^k),\tag{11.177}$$

于是 z^{k+1} 可很方便地求得: 首先解

$$L_ky^k=-\nabla f(x^k),\tag{11.178}$$

即 $y_1^k=-\partial f(x^k)/\partial x_1$ 和

$$y_i^k=-\frac{\partial f(x^k)}{\partial x_i}-\sum_{j=1}^{i-1}l_{ij}^k y_j^k,\quad i=2,\ \cdots,\ n.\tag{11.179}$$

再解

$$L_k^Tz^{k+1}=D_k^{-1}y^k,\tag{11.180}$$

即得 $z_n^{k+1}=y_n^k/d_{nn}^k$ 和

$$z_i^{k+1}=\frac{y_i^k}{d_{ii}^k}-\sum_{j=i+1}^{n} l_{ji}^k z_j^{k+1}, \quad i=n-1, \cdots, 1. \tag{11.181}$$

现在，如果利用关于 H_k 的对称秩 2 修改公式或是对称秩 1 形式，则可证明，$H_k^{-1}=B_k$ 的相应修改公式为

$$B_k=B_{k-1}+\mu_1^k y^k(y^k)^T+\mu_2^k w^k(w^k)^T, \tag{11.182}$$

其中 μ_1^k、μ_2^k 是数量，y^k、w^k 为向量，它们可以从 H_k 的具体修改公式，例如 DFP 和 BFS 公式中找到. B_k 的值通过下列公式逐步算得:

$$\hat{L}_k\hat{D}_k\hat{L}_k^T=B_{k-1}+\mu_1^k y^k(y^k)^T, \tag{11.183}$$

$$B_k=L_kD_kL_k^T=\hat{L}_k\hat{D}_k\hat{L}_k^T+\mu_2^k w^k(w^k)^T. \tag{11.184}$$

例如，如果利用 BFS 公式，则

$$\mu_1^k=\frac{1}{(p^k)^T\gamma^k}, \quad y^k=\gamma^k, \tag{11.185}$$

$$\mu_2^k=\frac{1}{(z^k)^T\nabla f(x^{k-1})}, \quad w^k=\nabla f(x^{k-1}). \tag{11.186}$$

作为改进了的变尺度算法，Gill-Murray 方法的重要特性是矩阵 B_k 保证为正定的，因此 z^{k+1}是一个下降方向. 关于这种方法的进一步细节，读者可参看 Gill、Murray[21]，Gill、Murray、Pitfield[22] 和 Murray[34]. Gill-Murray 方法的另一有趣性质是，它可以很容易改成不用导数的变尺度法. 这种方法将在下一节中讨论.

11.3 不用导数的变尺度算法

本章所述的变尺度法，在许多较难的试验函数上得到成功的应用，于是开始了扩大它们适用性的研究. 把这些方法用于求那些导数不能求得、或者难于计算的函数的极小. 既然所有这些方法仅用到一阶导数，自然的推广是用有限差分近似来代替解析式导数. 关于导数的有限差分有两种常用公式: 向前差分为

$$\frac{\Delta f(x)}{\Delta x_j}=\frac{f(x+h_jI_j)-f(x)}{h_j}, \quad j=1, \cdots, n, \tag{11.187}$$

中心差分为

$$\frac{\Delta f(x)}{\Delta x_j}=\frac{f(x+h_jI_j)-f(x-h_jI_j)}{2h_j}, \quad j=1, \cdots, n, \quad (11.188)$$

其中 $\Delta f(x)/\Delta x_j$ 是偏导数 $\partial f(x)/\partial x_j$ 的近似，h_j 为一个小的数，I_j 表示单位阵的第 j 列. 注意，如果 $f(x)$ 已知，为获得 $\nabla f(x)$ 的近似，向前差分式需要再计算 n 个函数值，相比之下，通常更为精确的中心差分公式则需要再计算 $2n$ 个函数值.

首次用有限差分代替导数来修改变尺度法的是 Stewart[52]. 他的方法的基础在于，使得利用有限差分时所含有的两种通常冲突的误差互相平衡. 如果 h_j 取得太大，作为有限差分公式基础的 f 的一阶近似就会变得不适当（截断误差），然而 h_j 取得太小时，在仅具有有限精确度的计算机中，两个函数值可以变得不可区别，因而一取差就会出现有效数字相互抵消的情况. Stewart 导出了这两种误差的估计，并提议选取使这两种误差估计相等的 h_j 值. 这个计算要解一个三次多项式方程，在其中也出现 $\nabla^2 f(x)$ 的对角元素的近似值，这些数由对 DFP 公式算得的 H_k 求逆而算得，求逆时利用本章前面提到过的 Sherman-Morrison[29] 公式. 如果计算所得的截断误差大于某个预定的容许范围，Stewart 提议用中心差分公式；否则，在变尺度算法，例如 DFP 方法中用向前差分公式作为梯度的近似. 每次迭代的线性搜索仅用到函数值的计算，它类似于 Powell[41] 在他那个不用导数的共轭方向算法中所提出的方法. Lill[13,32] 对这个线性搜索方法作了少许修改. 他也比较了 DFP 算法的 Stewart 修改方案和不用导数的 Powell 方法，他的结论是前者需要较少次数的函数值计算，特别是在大批变量的情况下.

Gill、Murray[21] 注意到，估计 $\nabla^2 f(x)$ 的对角元素的 Stewart 方法，其结果可能使 h_j 的值选取得不好. 因此，他们提出用有限差分作梯度近似的另一种方法. 在 Stewart 的方法中，从一次迭代到下一次迭代时，沿着坐标轴方向的间隔 h_j 是变化的，与此相反，Gill-Murray 在算法开始时固定 h_j 的大小. h_j 通常取得与变量 $x_j(j=1, \cdots, n)$ 的某个假定的范围成正比，函数的极小值点已知

是在这范围中．因为向前差分公式(11.187)需要较少次数的函数值计算，所以先用它来估计梯度 ∇f．现在假设在第 k 次迭代时我们已有了 $\nabla f(x^k)$ 的估计值．先检验这个估计值的大小．如果向前差分近似产生

$$\left\{\sum_{j=1}^{n}\left(\frac{\Delta f(x^k)}{\Delta x_j}\right)^2\right\}^{1/2}<\hat{\alpha}, \qquad (11.189)$$

其中 $\hat{\alpha}$ 是一个预先指定的常数，则就拒绝这个梯度估计值，而用中心差分公式(11.188)计算出一个新的估计值．有了用向前差分公式或中心差分公式算得的 $\nabla f(x^k)$ 的一个近似以后，再用前述的 Cholesky 分解来计算 z^{k+1}，并用一种在[22]中概述的有效线性搜索来计算 α_{k+1}^*，以近似地达到 $f(x^k+\alpha_{k+1}z^{k+1})$的极小．令 $x^{k+1}=x^k+\alpha_{k+1}^*z^{k+1}$，检验这一步所取的大小．如果

$$\|x^{k+1}-x^k\|=\|p^{k+1}\|<\hat{\beta}, \qquad (11.190)$$

其中 $\hat{\beta}$ 也是预先指定的常数，而且所用的是向前差分公式，则拒绝方向 z^{k+1}，利用中心差分公式来近似 $\nabla f(x^k)$，以求得方向 z^{k+1} 的新值．一旦中心差分公式被采用，以后逐次迭代中就都用这公式，直到步子变得大了，也就是说，直到对某个 $i>k$ 成立

$$\|p^i\|\geqslant\hat{\beta}. \qquad (11.191)$$

数值经验已证明[22, 27]，用和不用解析导数这两种形式 Gill-Murray 算法是极其成功的，特别当连同 BFS 修改公式一起使用时更是如此．这个方法的实现还包含许多细节，在这里不能重复了，有兴趣的读者可参看[22]．

前述两种方法都是用有限差分去近似一阶导数，然后利用某个已有的变尺度修改公式．另外还有两种算法可认为是不需计算导数的拟 Newton 方法．第一种是由 Greenstadt 提出，后来又加以改进的[26, 27]．它建立在同一作者更早工作[25]的基础上．在[25]中，变尺度校正矩阵 ΔH_k 是由变分原理导出的，这变分原理就是在满足割线关系式的约束条件下，求某个矩阵模的极小值．这些推导已修改成适应于不能利用解析导数的情况．Greenstadt[27] 的数值试验指出，在大多数情况下，Gill-Murray-Pitfield 方法执行得

较好.第二种方法属于 Fiacco、McCormick[17]，它由 $n(n+1)/2$ 个线性搜索构成，以得到 n 变量函数的梯度和 Hesse 阵的近似值．一阶和二阶导数的估计值一经算得，就取 Newton 步．这个算法的一个重要性态是，对于二次函数来说近似值是精确的，可用有限次迭代求得精确的极小值点．在一般非线性函数情况下，在 Newton 步之后，所有以前的导数估计值都无用了，要重新建立新的近似值．

11.4 基于非二次函数的极小化方法

几乎所有有效的多维极小化算法都要利用求极小的那个函数的二次近似式．虽然这种近似在最优点的小邻域内通常是适用的，但对于离最优点较远的点可能并不如此．例如，保证矩阵 H_k 正定性的拟 Newton 法对于不是二次的非凸函数，在远离极小值点处的 Hesse 逆阵可能得出坏的近似．所以，试着去修改现有算法，或导出明确地考虑非二次函数的新算法，看来是合乎逻辑的．本节介绍两种这样的算法，它们已在很有限的规模上进行了数值试验，与现有的变尺度法相比，还是有希望的．我们从 Biggs[4,5] 的**主幂法**开始．在这个方法中，对现有的修改公式诸如 DFP、BFS 或秩 1 公式作了一些变化，使得求极小的函数 f 的一些非二次性质能反映出来．变尺度算法隐蔽地假定了，在由算法产生的相继两点之间，f 是二次的．另一方面，Biggs 方法假定，如果定义 ϕ 为

$$\phi(\alpha)=f(x+\alpha z), \tag{11.192}$$

其中 x 和 z 为给定的向量，则一维函数 ϕ 能适当地表示成

$$\phi(\alpha)=a|\alpha-\alpha^*|^p+b, \tag{11.193}$$

其中 $a>0$，$p>1$．这是一个具有极小值点 α^* 的凸函数，它由 $(\alpha)_p$ 的项所支配．为了确定 p 的值，我们按下面方式进行：设在某一点 α^0 处已有了函数值 $\phi(\alpha^0)$，一阶导数值 $\phi'(\alpha^0)$ 和二阶导数的一个近似值 $\tilde{\phi}''(\alpha^0)=\eta\phi''(\alpha^0)$，其中 $\eta>0$ 是某一未知常数．现在令

$$\alpha^1=\alpha^0-\theta\frac{\phi'(\alpha^0)}{\tilde{\phi}''(\alpha^0)}, \tag{11.194}$$

并假设 $\phi(\alpha^1)$和 $\phi'(\alpha^1)$值已知. Biggs[4] 证明了，用这些数据能算出 p 和 η 来. 定义

$$\xi=\frac{\phi(\alpha^0)-\phi(\alpha^1)}{\phi'(\alpha^0)(\alpha^1-\alpha^0)}, \tag{11.195}$$

$$\beta=\frac{\phi'(\alpha^1)}{\phi'(\alpha^0)}, \tag{11.196}$$

未知数 p 和 η 可由解下列一对非线性方程得到:

$$\frac{\eta(p-1)(1-\beta)}{\theta p}+\frac{\beta}{p}=\xi, \tag{11.197}$$

$$\left|1-\frac{\theta}{\eta(p-1)}\right|^p\left(\frac{\eta(p-1)}{\eta(p-1)-\theta}\right)=\beta. \tag{11.198}$$

确定了 p 和 η 的值后，[4] 中还证明 ϕ 在 α^1 处的二阶导数的一个好的估计值为

$$\tilde{\phi}''(\alpha^1)=\frac{1}{\eta^*}\frac{\phi'(\alpha^0)-\phi'(\alpha^1)}{(\alpha^0-\alpha^1)}, \tag{11.199}$$

其中

$$\eta^*=\frac{\eta}{\theta}\left[1-\frac{\theta}{\eta(p-1)}\right]\left(\frac{1}{\beta}-1\right). \tag{11.200}$$

从下面的推导可见，变尺度法假定了 $\eta^*=1$. 假设利用BFS修正公式(11.170)，则 $B_k=H_k^{-1}$ 为

$$B_k=B_{k-1}+\frac{\gamma^k(\gamma^k)^T}{(p^k)^T\gamma^k}-\frac{B_{k-1}p^k(p^k)^TB_{k-1}}{(p^k)^TB_{k-1}p^k}. \tag{11.201}$$

因为 B_k 是求极小的函数 f 的 Hesse 阵 $\nabla^2 f(x)$的近似，f 沿 z^k 的二阶方向导数为$(\hat{z}^k)^TB_k\hat{z}^k$，其中 $\hat{z}^k=z^k/\|z^k\|$. 从(11.201)可得

$$(\hat{z}^k)^TB_k\hat{z}^k=\frac{[(\hat{z}^k)^T\gamma^k]^2}{(p^k)^T\gamma^k}=\frac{(\hat{z}^k)[\nabla f(x^k)-\nabla f(x^{k-1})]}{\|p^k\|}. \tag{11.202}$$

再从(11.192)和(11.202)推知，BFS 方法是在(11.199)中假定了 $\eta^*=1$. 读者可用 H_k 的 DFP 修改公式代替 (11.170) 而验证同样结果. 如果假定 f 沿 z^k 方向是非二次的，则二阶导数的适当表达式为

$$(\hat{z}^k)^T B_k \hat{z}^k = \frac{(\hat{z}^k)^T[\nabla f(x^k) - \nabla f(x^{k-1})]}{\eta^* \| p^k \|}, \tag{11.203}$$

容易验证，改变了的 BFS 公式

$$H_k = H_{k-1} + \left[\eta^* + \frac{(\gamma^k)^T H_{k-1}\gamma^k}{(p^k)^T\gamma^k}\right]\frac{p^k(p^k)^T}{(p^k)^T\gamma^k} - \frac{p^k(\gamma^k)^T H_{k-1}}{(p^k)^T\gamma^k} - \frac{H_{k-1}\gamma^k(p^k)^T}{(p^k)^T\gamma^k} \tag{11.204}$$

确实产生(11.203)．类似地，改变了的 DFP 公式为

$$H_k = H_{k-1} + \eta^* \frac{p^k(p^k)^T}{(p^k)^T\gamma^k} - \frac{H_{k-1}\gamma^k(\gamma^k)^T H_{k-1}}{(\gamma^k)^T H_{k-1}\gamma^k}. \tag{11.205}$$

我们已看到，η^* 的计算需要解两个非线性方程，因此计算略为长些．主幂法的一种简化形式是假定函数 f 沿一直线可表示成一个次数至多为三次的多项式；即可写成

$$\phi(\alpha) = a + b\alpha + c(\alpha)^2 + d(\alpha)^3. \tag{11.206}$$

Biggs[5] 用一些简单的代数运算证明了，在这种情形下 η^* 的显式表示为

$$\eta^* = \frac{\bar{\beta} - 1}{2(2\bar{\beta} + 1 - 3\bar{\xi})}, \tag{11.207}$$

其中 $\bar{\xi}$ 和 $\bar{\beta}$ 是待定的参数．具有改变了的修正公式的完整算法非常相似于任何其他拟 Newton 法．它不含有精确的线性搜索，而仅要求函数值的充分减小．除非后来的搜索方向 z^{k-1} 和 z^k 几乎平行，步长 $\bar{\alpha}_k$ 取法为

$$\bar{\alpha}_k = \begin{cases} \min\left[0.1, \dfrac{1}{\|z^k\|}\right], & k \leqslant n, \\ 1, & k > n. \end{cases} \tag{11.208}$$

令

$$x^k = x^{k-1} + \bar{\alpha}_k z^k, \tag{11.209}$$

并规定

$$\bar{\xi} = \frac{f(x^k) - f(x^{k-1})}{(p^k)^T \nabla f(x^{k-1})}, \tag{11.210}$$

$$\bar{\beta} = \frac{(p^k)^T \nabla f(x^k)}{(p^k)^T \nabla f(x^{k-1})}. \tag{11.211}$$

注意，如果求极小的函数 f 是二次的，即在 (11.206) 中 $d=0$，则 (11.207) 将给出 $\eta^*=1$，从而改进的方法将归结为普通的拟 Newton 算法。

下一个算法是建立在二次函数的某种推广之上的. 给定二次函数 f，它的唯一极小值点在 $x^*\in R^n$ 处. 我们可把 f 重写成

$$f(x)=\frac{1}{2}(x-x^*)^T Q(x-x^*)+f(x^*). \tag{11.212}$$

这个函数也满足

$$f(x)=\frac{1}{2}(x-x^*)^T \nabla f(x)+f(x^*). \tag{11.213}$$

我们可以把这个关系式推广. 如果

$$f(x)=\frac{1}{m}(x-x^*)^T \nabla f(x)+f(x^*), \tag{11.214}$$

其中 m 为一非零数，则定义 f 为 x^*-齐 m 次可微实函数. 齐 m 次函数的 Euler 关系式[3] 是 (11.214) 在 $x^*=f(x^*)=0$ 时的特殊情形. 换一种说法，使得关系式

$$(t)^m[f(x)-f(x^*)]=f(tx+(1-t)x^*)-f(x^*) \tag{11.215}$$

成立的可微函数 f 是满足(11.214)的. 我们知道，对于二次函数，Newton 法只要一步就找到精确的极小值点. 我们将看到这个结果也可推广，这借助于证明：步长有适当限定的 Newton 法恰好一步就可找到 x^*-齐次函数的精确极小值点. 对 (11.214) 式取微分得

$$\nabla f(x)=\frac{1}{m}\nabla^2 f(x)(x-x^*)+\frac{1}{m}\nabla f(x). \tag{11.216}$$

因此

$$x^*=x-(m-1)[\nabla^2 f(x)]^{-1}\nabla f(x). \tag{11.217}$$

例 11.4.1

考虑 $\bar{x}$-齐 4 次函数

$$f(x)=\left[\frac{1}{2}(x-\bar{x})^T Q(x-\bar{x})\right]^2, \tag{11.218}$$

其中

$$\bar{x}=(1,\ 1)^T,\quad Q=\begin{bmatrix} 3 & -1 \\ -1 & 1 \end{bmatrix}. \tag{11.219}$$

我们有

$$\nabla f(x)=[(x-\bar{x})^TQ(x-\bar{x})]Q(x-\bar{x}), \tag{11.220}$$

$$\nabla^2 f(x)=[(x-\bar{x})Q(x-\bar{x})]Q+2Q(x-\bar{x})(x-\bar{x})^TQ. \tag{11.221}$$

取 $x^0=(-2,\ 4)^T$，则 $\nabla f(x^0)=(-648,\ 324)^T$，

$$\nabla^2 f(x^0)=\begin{bmatrix} 450 & -198 \\ -198 & 126 \end{bmatrix},$$

$$[\nabla^2 f(x^0)]^{-1}=\frac{1}{972}\begin{bmatrix} 7 & 11 \\ 11 & 25 \end{bmatrix}. \tag{11.222}$$

因此

$$x^*=\begin{pmatrix} -2 \\ 4 \end{pmatrix}-\frac{3}{972}\begin{bmatrix} 7 & 11 \\ 11 & 25 \end{bmatrix}\begin{pmatrix} -648 \\ 324 \end{pmatrix}=\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tag{11.223}$$

正如所断定的那样，$x^*=\bar{x}$. ∎

Jacobson、Oksman[37] 发展了一种类似于变尺度法的算法，不用计算二阶导数，应用于求 x^*-齐次函数的极小时可在有限步内收敛. 整理(11.214)得到

$$x^T\nabla f(x)=(x^*)^T\nabla f(x)+mf(x)-mf(x^*). \tag{11.224}$$

定义 $w\in R$、$y\in R^{n+2}$、$\lambda\in R^{n+2}$ 如下：

$$w=x^T\nabla f(x),\quad \lambda=\begin{pmatrix} x^* \\ m \\ mf(x^*) \end{pmatrix},\quad y=\begin{pmatrix} \nabla f(x) \\ f(x) \\ -1 \end{pmatrix}, \tag{11.225}$$

在某点 $x^k\in R^n$ 处求值，我们可以把(11.224)写成

$$w^k=\lambda^Ty^k. \tag{11.226}$$

假设我们计算 $w^1,\ \cdots,\ w^{n+2}$ 和 $y^1,\ \cdots,\ y^{n+2}$ 使得 y^k 线性无关，则令

$$Y=\begin{bmatrix}(y^1)^T\\ \vdots\\ (y^{n+2})^T\end{bmatrix},\quad v=\begin{bmatrix}w^1\\ \vdots\\ w^{n+2}\end{bmatrix}. \tag{11.227}$$

(11.226)可写成

$$\lambda=Y^{-1}v, \tag{11.228}$$

其中"最优性向量"$\lambda\in R^{n+2}$包含着这个问题之解的全部信息.

已经建议,对矩阵Y^{-1}进行修改来递推地求解(11.228). 假设第k次迭代时有

$$\lambda^k=R_kv^k,\quad k=0,1,\cdots, \tag{11.229}$$

其中$R_k=Y_k^{-1}$,且$\lambda^k=[x^{*k}, m^k, m^kf(x^{*k})]^T$是最优性向量的第$k$次估计值. 取$R_0=I$,且$v^0=\lambda^0$为一给定的$n+2$维向量. 在每次迭代时,用算得的向量$y^k$逐个更换$R_0^{-1}$的行向量,用算得的数$w^k=(\lambda^{k-1})^Ty^k$逐个更换$v^0$的元素. 即

$$R_k^{-1}=R_{k-1}^{-1}+I_j[(y^k)^T-I_j^TR_{k-1}^{-1}], \tag{11.230}$$

$$v^k=v^{k-1}+I_j[(w^k)^T-I_j^Tv^{k-1}], \tag{11.231}$$

其中I_j是$(n+2)\times(n+2)$单位阵的第j列,且$j=k$. 现在由秩1求逆公式[29]有

$$R_k=R_{k-1}-\frac{R_{k-1}I_j[(y^k)^TR_{k-1}-I_j^T]}{(y^k)^TR_{k-1}I_j}. \tag{11.232}$$

从(11.229), (11.231)和(11.232)得到λ^k的修改公式

$$\lambda^k=\lambda^{k-1}+\frac{R_{k-1}I_j[w^k-(y^k)^T\lambda^{k-1}]}{(y^k)^TR_{k-1}I_j}. \tag{11.233}$$

对n个变量的x^*-齐次函数, 向量λ^{n+2}等于真正的最优性向量λ. 因此$n+2$步后Jacobson-Oksman算法(参看下面)求得了它的精确的极小值点,齐次性的次数以及求极小的函数的最优值.

在概念化的形式下,完整的算法由下列步骤组成:

1. 假定$x^0\in R^n$给定,置$m^0=2$, $m^0f(x^{*0})=0$(指(11.225)中λ的相应初值——译者), $k=0$,对所有的x,令η为$f(x)$值的某一下界.

2. 令

$$x^1 = x^0 - \theta_0 \nabla f(x^0), \tag{11.234}$$

其中

$$\theta_0 = \min\left\{1, \frac{m^0[\eta - f(x^0)]}{(-\nabla f(x^0))^T \nabla f(x^0)}\right\}. \tag{11.235}$$

如果 $f(x^0) > f(x^1)$，转向步骤 3；否则用某种线性搜索找一个 θ_0，使得 $f(x^0) > f(x^1)$ 再转向步骤 3.

3. 令 $\lambda^0 = [x^1, 2, 0]^T$；$R_0 = I$ 和 $j = 1$.

4. 计算 y^{k+1}, w^{k+1}. 利用(11.232)和 (11.233) 计算 R_{k+1} 和 λ^{k+1}.

5. 改变 k 为 $k+1$. 如果 $j = n+2$, 重新令 $j = 1$；否则改变 j 为 $j+1$.

6. 令(下式中 x^{*k} 由 λ^k 中得到, 见(11.225)——译者)

$$x^{k+1} = x^k + \theta_k \sigma_k (x^k - x^{*k}), \tag{11.236}$$

其中 σ_k 或者为 $+1$ 或者为 -1，它的符号与 $-(x^k - x^{*k})^T \nabla f(x^k)$ 的符号一致, θ_k 给定为

$$\theta_k = \min\left\{1, \frac{m^k[\eta - f(x^k)]}{\sigma_k (x^k - x^{*k})^T \nabla f(x^k)}\right\}. \tag{11.237}$$

这里我们假定 $m^k > 0$. 此外, 如果 $m^k \leqslant 0$, 前面公式中就用 $m^k = 2$. 如果由(11.237)得出的 θ_k 值使得 $f(x^{k+1}) < f(x^k)$, 则转向步骤 4；否则再用某个线性搜索使得函数值充分减小，然后再转向步骤 4.

当 $\|\nabla f(x^{k+1})\|$ 小于某一小的正数时，算法终止. 贯穿整个算法, 在(11.232)和(11.233)中出现的分母都要检验. 如果

$$|(y^{k+1})^T R_k I_j| < \varepsilon, \tag{11.238}$$

其中 ε 是一小的正数, 就令 x^{k+1} 为新的 x^0, 而回到步骤 1.

例 11.4.2

我们用上例中两个变量的 x^*-齐 4 次函数来说明 Jacobson-Oksman 算法.

假设我们仍从 $x^0 = (-2, 4)^T$ 出发, $f(x^0) = 729$. 令 $\eta = 0$. 现在

$$\theta_0=\min\left\{1,\left[\frac{2(-729)}{-(648)^2-(324)^2}\right]\right\} \tag{11.239}$$

$$=\min\{1,\ 0.00278\}=0.00278. \tag{11.240}$$

因此

$$x^1=\begin{pmatrix}-0.2\\3.1\end{pmatrix},$$

$$f(x^1)=47.4032, \tag{11.241}$$

$$\nabla f(x^1)=\begin{pmatrix}-78.489\\45.441\end{pmatrix}.$$

其次令

$$\lambda^0=\begin{pmatrix}-0.2\\3.1\\2\\0\end{pmatrix},\quad R_0=I,\ j=1,\ k=0, \tag{11.242}$$

$$y^1=\begin{pmatrix}-78.4890\\45.4410\\47.4032\\-1\end{pmatrix},\quad w^1=156.565, \tag{11.243}$$

$$R_1=\begin{bmatrix}-0.0127 & 0.5789 & 0.6039 & -0.0127\\0 & 1 & 0 & 0\\0 & 0 & 1 & 0\\0 & 0 & 0 & 1\end{bmatrix},\ \lambda^1=\begin{pmatrix}1.008\\3.1\\2\\0\end{pmatrix}. \tag{11.244}$$

这样，预测的极小值点在 $x^{*1}=(1.008,3.1)^T$. 现在令 $k=1, j=2$, 且有 $(x^1-x^{*1})=(-1.208,\ 0)^T$.

下一个点为

$$x^2=\begin{pmatrix}-0.2\\3.1\end{pmatrix}+\theta_1\sigma_1\begin{pmatrix}-1.208\\0\end{pmatrix}. \tag{11.245}$$

因为 $(x^1-x^{*1})^T\nabla f(x^1)>0$, 可知 $\sigma_1=-1$,

$$\theta_1=\min\left\{1,\left[\frac{2(-47.4032)}{-0.94806}\right]\right\}=1, \tag{11.246}$$

所以 $x^2=(1.008,\ 3.1)^T$，在这点处

$$f(x^2)=4.7896<f(x^1). \tag{11.247}$$

这一点被接受并计算

$$\nabla f(x^2)=\begin{pmatrix}-9.0881\\ 9.1572\end{pmatrix}, \tag{11.248}$$

$$y^2=\begin{pmatrix}-9.0881\\ 9.1572\\ 4.7896\\ -1\end{pmatrix},\quad w^2=19.2275. \tag{11.249}$$

由此得

$$R_2=\begin{bmatrix}-0.0299 & 0.1486 & 0.7078 & 0.1187\\ -0.0297 & 0.2567 & 0.1795 & 0.2270\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix} \tag{11.250}$$

和 $\lambda^2=(-0.416,\ 0.641,\ 2,\ 0)^T$. 这样，下一个预测的极小值点在 $x^{*2}=(-0.416,\ 0.641)^T$. 令 $k=2$, $j=3$, 且有 $(x^2-x^{*2})=(1.424,\ 2.459)^T$.

下一个点为

$$x^3=\begin{pmatrix}1.008\\ 3.1\end{pmatrix}+\theta_2\sigma_2\begin{pmatrix}1.424\\ 2.459\end{pmatrix}. \tag{11.251}$$

因为 $(x^2-x^{*2})^T\nabla f(x^2)>0$，令 $\sigma_2=-1$，

$$\theta_2=\min\left\{1,\left[\frac{2(-4.7896)}{(-9.5792)}\right]\right\}=1. \tag{11.252}$$

因此 $x^3=(-0.416,\ 0.641)^T$，而

$$f(x^3)=6.5668>f(x^2). \tag{11.253}$$

即函数值增大了. 所以 x^3 这点被抛弃了，不妨以三次近似来寻找新的点 x^3. 这种方法(参看第 8 章)的一次迭代给出具有较小函数值的点

$$x^3=\begin{pmatrix} 0.3721 \\ 2.0017 \end{pmatrix},\ f(x^3)=2.9661,\ \nabla f(x^3)=\begin{pmatrix} -9.9392 \\ 5.6134 \end{pmatrix}. \tag{11.254}$$

因而,

$$y^3=\begin{pmatrix} -9.9392 \\ 5.6134 \\ 2.9661 \\ -1 \end{pmatrix},\quad w^3=7.5386, \tag{11.255}$$

$$R_3=\begin{bmatrix} 0.0003 & 0.1403 & -0.2312 & -0.0906 \\ -0.0221 & 0.2546 & -0.0586 & 0.1739 \\ 0.0427 & -0.0118 & -0.3266 & -0.2957 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{11.256}$$

和 $\lambda^3=(1.0,\ 1.0,\ 4.0,\ 0.0)^T$.

这样, 下一个预测的极小值点在 $x^{*3}=(1.0,\ 1.0)^T$. 令 $k=3$, $j=4$ 和 $(x^3-x^{*3})=(-0.6279,\ 1.0017)^T$. 那末

$$x^4=\begin{pmatrix} 0.3721 \\ 2.0017 \end{pmatrix}+\theta_3\sigma_3\begin{pmatrix} -0.6279 \\ 1.0017 \end{pmatrix}. \tag{11.257}$$

因为 $(x^3-x^{*3})^T\nabla f(x^3)>0$, 可得 $\sigma_3=-1$,

$$\theta_3=\min\left\{1,\left[\frac{4(-2.9661)}{-11.8644}\right]\right\}=1. \tag{11.258}$$

因此 $x^4=(1.0,\ 1.0)^T$, $f(x^4)=0.0$, $\nabla f(x^4)=(0.0,\ 0.0)^T$, 算法在精确极小值点处终结. ▌

练　习

11.A. 在 (11.25) 中取 $\omega=0$, 求出 Huang 族矩阵 H_k 的修改公式, 并使其满足

(a) ΔH_k 为秩 2 矩阵.

(b) ΔH_k 为对称秩 1 矩阵.

在 (11.25) 中取 $\omega=-1$, 求修改公式, 并使其满足

(c) ΔH_k 为对称秩 2 矩阵.

(d) ΔH_k 为对称秩 1 矩阵.

11.B. 应用建立在上题(a)和(c)的修改公式基础上的变尺度算法，对下列二次函数求极小：

$$f(x)=(x_1-x_2+x_3)^2+(-x_1+x_2+x_3)^2+(x_1+x_2-x_3)^2. \quad (11.259)$$

取初始点为 $x^0=\left(\frac{1}{2}, 1, \frac{1}{2}\right)^T$，$H_0=I$，采用精确的线性搜索．算法所产生的 x^2 和 x^3 是否一致？你是否获得二次终结性质？H_n 是否等于 Q^{-1}？

11.C. 对于求 R^n 上函数 f 的极小，最早的变尺度算法之一是由 Zoutendijk[55] 提出的，可描述如下：令 x^0 为 R^n 中任一点，取 $H_0=I$，

$$H_k=H_{k-1}-\frac{H_{k-1}\gamma^k(\gamma^k)^T H_{k-1}}{(\gamma^k)^T H_{k-1}\gamma^k}. \quad (11.260)$$

在每次迭代中都执行精确的线性搜索．证明由此修正公式所得的 H_k 是 Huang 族矩阵．ω 的值是什么？应用 Zoutendijk 方法求两个变量的二次函数的极小．H_2 是什么？

11.D. 设 $n\times n$ 阵 H_{k-1} 是对称正定的，且采用 DFP 公式进行修改．证明：如果 $(p^k)^T\gamma^k>0$，则 H_k 也是对称正定的．

11.E. 已经知道，DFP 方法对于求极小的函数的数乘运算不是不变的．设 x^1, x^2, …和 $\hat{x}^1$, $\hat{x}^2$, …分别为用 DFP 方法求函数 $f(x)$ 和函数 $\hat{f}(x)=cf(x)$ 的极小所产生的点列，其中 c 为一正数，$c\neq 1$．说明这两个点列并不恒同．你能否把这个结果推广到 Broyden 子族算法和 Huang 族算法上去？讨论：对 $x\in R^n$ 进行数乘，得 $ax(a\in R)$，代替求 $f(x)$ 的极小，我们求 $f(ax)$ 的极小，这对由 DFP 方法所得的点列产生什么影响？

11.F. Oren[38] 阐明了对 DFP 算法稍作改变的一种方式，使得它在上题所讨论的函数数乘或变量数乘下，都是不变的．例如，他提议矩阵 H_k 在修改之前乘以因子

$$\zeta_k=\frac{(p^k)^T\gamma^k}{(\gamma^k)^T H_{k-1}\gamma^k}. \quad (11.261)$$

这等价于采用下述改变了的 DFP 修改公式：

$$H_k=\left[H_{k-1}-\frac{H_{k-1}\gamma^k(\gamma^k)^T H_{k-1}}{(\gamma^k)^T H_{k-1}\gamma^k}\right]\frac{(p^k)^T\gamma^k}{(\gamma^k)^T H_{k-1}\gamma^k}+\frac{p^k(p^k)^T}{(p^k)^T\gamma^k}. \quad (11.262)$$

证明：建立在上述修改公式基础上的变尺度算法，在求 $f(x)$ 和

$\check{f}(\alpha x)$极小时，相应地产生相同的点列 $x^1, x^2, \cdots$和 $\hat{x}^1, \hat{x}^2, \cdots$.

11.G. 设用DFP方法求 R^n 上具有对称正定阵 Q 的二次函数的极小. 证明在这种情况下，修改公式可写成

$$H_k = P_k + C_k, \quad k=1, \cdots \tag{11.263}$$

其中

$$P_k = P_{k-1} - \frac{P_{k-1}\gamma^k(P_{k-1}\gamma^k)^T}{(\gamma^k)^T P_{k-1}\gamma^k}, \tag{11.264}$$

$$C_k = C_{k-1} + \frac{p^k(p^k)^T}{(p^k)^T\gamma^k}, \tag{11.265}$$

且 $P_0 = H_0$, $C_0 = 0$. [提示: $H_{k-1}\gamma^k = P_{k-1}\gamma^k$.] 把上述关系式结合(11.89)以证明 $C_n = Q^{-1}$. 是否 $P_n = 0$?

11.H. 证明: 如果(11.69)成立，则由(11.67)和(11.68)给出的Broyden子族矩阵 H_k 将产生 $g^{k+1}=0$.

11.I. 设 A 为 n 阶非异方阵，且 U、V 为 $n\times m$ 阵，$n\geqslant m$. 证明: $(A+UV^T)$ 有逆阵的充分必要条件为 $(I+V^TA^{-1}U)$ 有逆阵，其中 I 为 m 阶单位阵. 并证明

$$(A+UV^T)^{-1} = A^{-1} - A^{-1}U(I+V^TA^{-1}U)^{-1}V^TA^{-1}. \tag{11.266}$$

把这个公式同(11.139)式比较.

11.J. 证明: 如果一个奇异的 H_k 用于变尺度算法中，则所有后继矩阵也都是奇异的(除了非常特殊的情况). 奇异矩阵 H_k 对后继的搜索方向有什么影响?

11.K. 证明DFP和BFS修改公式之间的"互补"关系. 如果 H_k 采用对称秩1公式来修改，那么近似Hesse阵的修改公式是什么?

11.L. Powell[43] 提出了一种求 R^n 上函数 $f(x)$ 极小的变尺度方法，在这算法中，近似Hesse阵采用下列公式进行修改:

$$B_k = B_{k-1} + \frac{(\gamma^k - B_{k-1}p^k)(p^k)^T + p^k(\gamma^k - B_{k-1}p^k)^T}{(p^k)^Tp^k} - \frac{(p^k)^T(\gamma^k - B_{k-1}p^k)p^k(p^k)^T}{[(p^k)^Tp^k]^2}. \tag{11.267}$$

(a) 寻找Hesse逆阵的近似矩阵 $H_k = B_k^{-1}$ 的修改公式，并证明 H_k 满足割线关系式.

(b) 设 f 是具有非异阵 Q 的二次函数. 证明

$$B_k - Q = \left[I - \frac{p^k(p^k)^T}{(p^k)^Tp^k}\right](B_{k-1}-Q)\left[I - \frac{p^k(p^k)^T}{(p^k)^Tp^k}\right]. \tag{11.268}$$

(c) 证明: 如果修改公式(11.267)重复应用 n 次，且向量 $p^1, p^2, \cdots,$

p^n 相互正交，则在二次函数情况下，$B_n=Q$.

11.M. 证明：建立在对称秩1公式(11.158)基础上的 Murtagh-Sargent 变尺度算法的步长 α_k，如果是选得使具有正定阵 Q 的二次函数沿 s^k 方向达到极小，则搜索方向 $s^1, \cdots, s^n$ 关于 Q 相互共轭。证明时不要利用 H_k 是 Huang 族矩阵这一事实.

11.N. 假设 R^n 上连续可微函数的拟 Newton 极小化算法由下述一般公式给定：

$$x^{k+1}=x^k+\alpha_{k+1}s^{k+1}, \quad k=0, 1, \cdots \tag{11.269}$$

其中

$$s^{k+1}=-H_k\nabla f(x^k), \quad k=0, 1, \cdots \tag{11.270}$$

H_k 是对称正定阵，建议沿 s^{k+1} 方向作 f 的单个二次近似式，以此来选取步长，做法如下：

令

$$x^r=x^k+\alpha_r s^{k+1}, \quad \alpha_r>0, \tag{11.271}$$

x^r 是使得 $f(x^r)>f(x^k)$ 的任意点，选取

$$\alpha_{k+1}=\frac{1}{2}\frac{(\alpha_r)^2(\nabla f(x^k))^T H_k\nabla f(x^k)}{(f(x^r)-f(x^k))+\alpha_r(\nabla f(x^k))^T H_k\nabla f(x^k)}. \tag{11.272}$$

证明：x^{k+1} 是 f 沿 s^{k+1} 方向的二次近似式的极小值点，并且 $0<\alpha_{k+1}<\alpha_r$.

11.O. 对本章中提到的基于矩阵修改公式的变尺度算法，从下列理论性质进行分析：

(a) 二次终结性质(有或没有精确线性搜索).

(b) 矩阵 H_k 的正定性.

(c) 稳定性.

(d) 矩阵 H_k 趋于$[\nabla^2 f(x^*)]^{-1}$的收敛性.

如果需要，可查阅适当的参考文献.

参 考 文 献

1. BARD, Y., "On a Numerical Instability of Davidon-Like methods," *Math. Comp.*, **22**, 665–666 (1968).
2. BARNES, J. G. P., "An Algorithm for Solving Non-Linear Equations Based on the Secant Method," *Computer J.*, **8**, 66–72 (1965).
3. BARTLE, R. G., *The Elements of Real Analysis*, 2nd ed., John Wiley & Sons, New York, 1976.

4. BIGGS, M. C., "Minimization Algorithms Making Use of Non-quadratic Properties of the Objective Function," *J. Inst. Math. Appl.*, **8**, 315–327 (1971).

5 BIGGS, M. C., "A Note on Minimization Algorithms which Make Use of Non-Quadratic Properties of the Objective Function," *J. Inst. Math. Appl.*, **12**, 337–338 (1973).

6. BROYDEN, C. G., "A Class of Methods for Solving Nonlinear Simultaneous Equations," *Math. Comp.*, **19**, 577–593 (1965).

7. BROYDEN, C. G., "Quasi-Newton Methods and their Application to Function Minimisation," *Math. Comp.*, **21**, 368–381 (1967).

8. BROYDEN, C. G., "The Convergence of a Class of Double-Rank Minimization Algorithms 1. General Considerations," *J. Inst. Math. Appl.*, **6**, 76–90 (1970).

9. BROYDEN, C. G., "The Convergence of a Class of Double-Rank Minimization Algorithms 2. The New Algorithm," *J. Inst. Math. Appl.*, **6**, 222–231 (1970).

10. BROYDEN, C. G., "The Convergence of Single-Rank Quasi-Newton Methods," *Math. Comp.*, **24**, 365–382 (1970).

11. DAVIDON, W. C., "Variable Metric Method for Minimization," AEC Research and Development Report ANL-5990 (Rev.), November 1959.

12. DAVIDON, W. C., "Variance Algorithm for Minimization," *Computer J.*, **10**, 406–410 (1968).

13. DENNIS, J. E., "On the Convergence of Broyden's Method for Nonlinear Systems of Equations," *Math. Comp.*, **25**, 559–567 (1971).

14. DIXON, L. C. W., "Quasi-Newton Algorithms Generate Identical Points," *Math. Prog.*, **2**, 383–387 (1972).

15. DIXON, L. C. W., "Quasi-Newton Techniques Generate Identical Points. II. The Proofs of Four New Theorems," *Math. Prog.*, **3**, 345–358 (1972).

16. DIXON, L. C. W., "The Choice of Step Length a Crucial Factor in the Performance of Variable Metric Algorithms," in *Numerical Methods for Non-Linear Optimization*, F. Lootsma (Ed.), Academic Press, London, 1972.

17. FIACCO, A. V., and G. P. MCCORMICK, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, John Wiley & Sons, New York, 1968.

18. FLETCHER, R., "A Review of Methods for Unconstrained Optimization," in *Optimization*, R. Fletcher (Ed.), Academic Press, London, 1969.

19. FLETCHER, R., "A New Approach to Variable Metric Algorithms," *Computer J.*, **13**, 317–322 (1970).

20. FLETCHER, R., and M. J. D. POWELL, "A Rapidly Convergent Descent Method for Minimization," *Computer J.*, **6**, 163–168 (1963).

21. GILL, P. E., and W. MURRAY, "Quasi-Newton Methods for Unconstrained Optimization," *J. Inst. Math. Appl.*, **9**, 91–108 (1972).

22. GILL, P. E., W. MURRAY, and R. A. PITFIELD, "The Implementation of Two Revised Quasi-Newton Algorithms for Unconstrained Optimization," National Physical Laboratory DNAC Report No. 11, April 1972.

23. Goldfarb, D., "Sufficient Conditions for the Convergence of a Variable Metric Algorithm," in *Optimization*, R. Fletcher (Ed.), Academic Press, London, 1969.

24. Goldstein, A. A., and J. F. Price, "An Effective Algorithm for Minimization," *Numerische Mathematik*, **10**, 184–189 (1967).

25. Greenstadt, J., "Variations on Variable-Metric Methods," *Math. Comp.*, **24**, 1–22 (1970).

26. Greenstadt, J., "A Quasi-Newton Method with No Derivatives," *Math. Comp.*, **26**, 145–166 (1972).

27. Greenstadt, J., "Improvements in a QNWD Method," IBM Palo Alto Scientific Center Tech. Report No. 320-3306, October 1972.

28. Hestenes, M. R., "Multiplier and Gradient Methods," *J. Optimization Theory & Appl.*, **4**, 303–320 (1969).

29. Householder, A. S., *The Theory of Matrices in Numerical Analysis*, Blaisdell Publishing Co., New York, 1964.

30. Huang, H. Y., "Unified Approach to Quadratically Convergent Algorithms for Function Minimization," *J. Optimization Theory & Appl.*, **5**, 405–423 (1970).

31. Jacobson, D. H., and W. Oksman, "An Algorithm that Minimizes Homogeneous Functions of N Variables in $N + 2$ Iterations and Rapidly Minimizes General Functions," *J. Math. Anal. & Appl.*, **38**, 535–552 (1972).

32. Lill, S. A., "A Modified Davidon Method for Finding the Minimum of a Function Using Difference Approximation for Derivatives," *Computer J.*, **13**, 111–113 (1970).

33. Lootsma, F. A. (Ed.), *Numerical Methods for Non-Linear Optimization*, Academic Press, London, 1972.

34. Murray, W., "Failure, the Causes and Cures," in *Numerical Methods for Unconstrained Optimization*, W. Murray (Ed.), Academic Press, London, 1972.

35. Murray, W., "The Relationship between the Approximate Hessian Matrices Generated by a Class of Quasi-Newton Methods," National Physical Laboratory DNAC Report No. 12, January 1972.

36. Murtagh, B. A., and R. W. H. Sargent, "A Constrained Minimization Method with Quadratic Convergence," in *Optimization*, R. Fletcher (Ed.), Academic Press, London, 1969.

37. Murtagh, B. A., and R. W. H. Sargent, "Computational Experience with Quadratically Convergent Minimisation Methods," *Computer J.*, 13, 185–194 (1970).

38. Oren, S. S., "On the Selection of Parameters in Self-Scaling Variable Metric Algorithms," *Math. Prog.*, 7, 351–367 (1974).

39. Ortega, J. M., and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

40. PEARSON, J. D., "Variable Metric Methods of Minimisation," *Computer J.*, **12**, 171–178 (1969).

41. POWELL, M. J. D., "An Efficient Method for Finding the Minimum of a Function of Several Variables without Calculating Derivatives," *Computer J.*, **7**, 155–162 (1964).

42. POWELL, M. J. D., "Rank One Methods for Unconstrained Optimization," in *Integer and Nonlinear Programming*, J. Abadie (Ed.), North-Holland Publishing Co., Amsterdam, 1970.

43. POWELL, M. J. D., "A New Algorithm for Unconstrained Optimization," in *Nonlinear Programming*, J. B. Rosen, O. L. Mangasarian, and K. Ritter (Eds.), Academic Press, New York, 1970.

44. POWELL, M. J. D., "On the Convergence of Variable Metric Algorithms," *J. Inst. Math. Appl.*, **7**, 21–36 (1970).

45. POWELL, M. J. D., "Some Properties of the Variable Metric Algorithm," in *Numerical Methods for Non-Linear Optimization*, F. Lootsma (Ed.), Academic Press, London, 1972.

46. POWELL, M. J. D., "Quadratic Termination Properties of Minimization Algorithms I. Statement and Discussion of Results," *J. Inst. Math. Appl.*, **10**, 333–342 (1972).

47. POWELL, M. J. D., "Quadratic Termination Properties of Minimization Algorithms II. Proofs of Theorems," *J. Inst. Math. Appl.*, **10**, 343–357 (1972).

48. SARGENT, R. W. H., "Numerical Optimization Techniques," (mimeographed lecture notes), Dept. of Chemical Engineering and Chemical Technology, Imperial College, London, 1972.

49. SHANNO, D. F., "Conditioning of Quasi-Newton Methods for Function Minimization," *Math. Comp.*, **24**, 647–656 (1970).

50. SHANNO, D. F., and P. C. KETTLER, "Optimal Conditioning of Quasi-Newton Methods," *Math. Comp.*, **24**, 657–664 (1970).

51. SORENSON, H. W., "Comparison of Some Conjugate Direction Procedures for Function Minimization," *J. Franklin Inst.*, **288**, 421–441 (1969).

52. STEWART, G. W., "A Modification of Davidon's Minimization Method to Accept Difference Approximations of Derivatives," *J. Assoc. Comp. Mach.*, **14**, 72–83 (1967).

53. WILKINSON, J. H., *The Algebraic Eigenvalue Problem*, Oxford University Press, London, 1965.

54. WOLFE, P., "The Secant Method for Simultaneous Non-linear Equations," *Comm. Assoc. Comp. Mach.*, **2**, 12–13 (1959).

55. ZOUTENDIJK, G., *Methods of Feasible Directions*, Elsevier Publishing Co., Amsterdam, 1960.

第 12 章

惩罚函数法

本章开始讨论求解约束非线性规划的方法. 我们主要处理如下问题: 如何把约束规划变换为一个或多个等价的无约束规划, 使之能运用前几章讨论的方法求解.

在所有惩罚函数法背后的直观想法是很简单的. 假设我们要在 R^n 的一个真子集 X 上寻求一个实值函数的极小值点. 当然, 如同我们即将看到的那样, 对目标函数作某些修改后, 这个约束最优化问题能变换为无约束最优化问题. 定义

$$P(x)=\begin{cases}0, & x\in X,\\ +\infty, & x\overline{\in} X,\end{cases} \tag{12.1}$$

并考虑下式给出的**增广目标函数** F 的无约束极小化:

$$\min_{x\in R^n} F(x)=f(x)+P(x), \tag{12.2}$$

其中 f 假定在 R^n 上定义. 点 x^* 使 F 极小的充要条件是它也使 f 在 X 上极小. 函数 P 称为**惩罚函数**, 它对落在能行集外的点加以(无限值)惩罚. 然而, 在实际中, 因为 F 在 X 的边界上不连续、在 X 外取无限值, 所以无约束最优化(12.2)无法实现(某些平凡的情况可能除外). 用某个"大的"有限数惩罚代替 $+\infty$, 并不能使问题简化, 因为数值计算的困难仍然保留. 而且, 不作附加假定的话, 增广的处处有限的目标函数的极小与 f 在 X 上的极小可能不重合.

用惩罚函数求解约束问题的最早想法, 包含一个无约束极小化的序列, 在这个序列中, 惩罚参数从一个极小化转到另一个极小化时被调节, 使得无约束极小值点列收敛于约束问题的一个能行点, 它满足某些必要或充分的最优性条件. 通过有关的无约束问题来研究约束问题的最初建议大概是属于 Courant[9] 和

Frisch[22] 的.

前三节处理这样一些方法，它们之间的差别在于趋近于解的途径和参数的选择. 这些方法已成功地在若干应用中被实现（例如，参看[4, 21, 31, 51]），而且被认为是目前求解约束问题的主要的有效工具之一.

近年来的许多研究工作，致力于通过免除求解一系列无约束问题，以改进惩罚函数法. 导出了约束问题的某些变换，在其中，只要取充分大的参数，就只需一次无约束极小化. 这样的变换与前几章中分析过的非线性规划的 Lagrange 式以及最优条件有密切联系. 这些方法在 12.4 节和 12.5 节中讲述. 最后，在末了一节讨论惩罚函数法的某些计算背景.

12.1 外部惩罚函数

本节提出一类求解基本非线性规划的方法，该规划在第 3 章首次引入：

(P) $$\min f(x) \tag{12.3}$$

受限制于约束

$$g_i(x) \geqslant 0, \quad i=1, \cdots, m, \tag{12.4}$$

$$h_j(x)=0, \quad j=1, \cdots, p, \tag{12.5}$$

其中 $f, g_1, \cdots, g_m, h_1, \cdots, h_p$ 假定在 R^n 上连续. 令 X 表示能行集，即

$$X=\{x: x \in R^n,\ g_i(x) \geqslant 0,\ i=1, \cdots, m;\ h_j(x)=0,\ j=1, \cdots, p\}. \tag{12.6}$$

外部惩罚函数法通常用这样一系列无约束极小化问题求解(P)；这一系列问题的最优解从能行集外部趋于 (P) 的解. 在这一系列无约束最优化中，惩罚加于每个 $x \notin X$，随着由一个问题进入另一个问题，惩罚逐步增加，促使无约束最优点走向能行集. 这里提出的算法应归于 Zangwill[52]，那里能找到更详细的内容.

定义变量 $\eta \in R$ 的实值连续函数 ψ 和 ζ 如下：

$$\psi(\eta)=|\min(0, \eta)|^{\alpha}, \tag{12.7}$$

$$\zeta(\eta)=|\eta|^{\beta}, \tag{12.8}$$

其中 $\alpha\geqslant1$ 和 $\beta\geqslant1$ 是给定的常数，通常等于 1 或 2．令

$$s(x)=\sum_{i=1}^{m}\psi(g_i(x))+\sum_{j=1}^{p}\zeta(h_j(x)), \tag{12.9}$$

或

$$s(x)=\sum_{i=1}^{m}|\min[0,\ g_i(x)]|^{\alpha}+\sum_{j=1}^{p}|h_j(x)|^{\beta}. \tag{12.10}$$

这连续函数称为问题(P)的**损失函数**．注意

$$s(x)=0,\quad x\in X, \tag{12.11}$$

$$s(x)>0,\quad x\overline{\in}X. \tag{12.12}$$

对于任何正数 ρ，我们定义问题(P)的增广目标函数为

$$F(x,\ \rho)=f(x)+\frac{1}{\rho}s(x), \tag{12.13}$$

并且注意到，$F(x,\ \rho)=f(x)$ 成立的充要条件为 x 是能行的，否则 $F(x,\ \rho)>f(x)$．当 $\rho\to0$ 时 $s(x)/\rho$ 这一项近似于 (12.1) 中不连续的惩罚函数 $P(x)$．外部惩罚函数法由求解如下一系列无约束最优化问题组成：对于 $k=0,\ 1,\ 2,\ \cdots$，

$$(\mathrm{EP}^k)\qquad \min_{x\in R^n}F(x,\ \rho^k)=f(x)+\frac{1}{\rho^k}\left\{\sum_{i=1}^{m}|\min[0,\ g_i(x)]|^{\alpha}+\sum_{j=1}^{p}|h_j(x)|^{\beta}\right\}, \tag{12.14}$$

其中 ρ^k 是一严格递减的正数列．记 x^{k*} 为(EP^k)的最优解，构造点列 $\{x^{k*}\}$，在(P)的较为温和的条件下，它有一个子列收敛于(P)的一个最优解．

外部惩罚函数的上述推导能够推广．令 r 是变量 $\rho\in R$ 的一个连续实值函数，使得 $\rho^1>\rho^2>0$ 蕴涵 $r(\rho^2)>r(\rho^1)>0$，并且对每一具有性质

$$\lim_{k\to\infty}\{\rho^k\}=0 \tag{12.15}$$

的严格递减正数列 $\{\rho^k\}$，有

$$\lim_{k\to\infty}\{r(\rho^k)\}=+\infty. \tag{12.16}$$

令 s 也是满足 (12.11) 和 (12.12) 的任意连续函数，则

$r(\rho^k)s(x)$是一个**外部惩罚函数**，并且

$$(\mathrm{EP}^k) \qquad \min_x F(x, \rho^k) = f(x) + r(\rho^k)s(x) \tag{12.17}$$

是相应的无约束最优化问题. 在图 12.1 中，我们用一个简单的情形来说明这个想法，其中能行集 X 是实轴上的闭区间$[a, b]$.

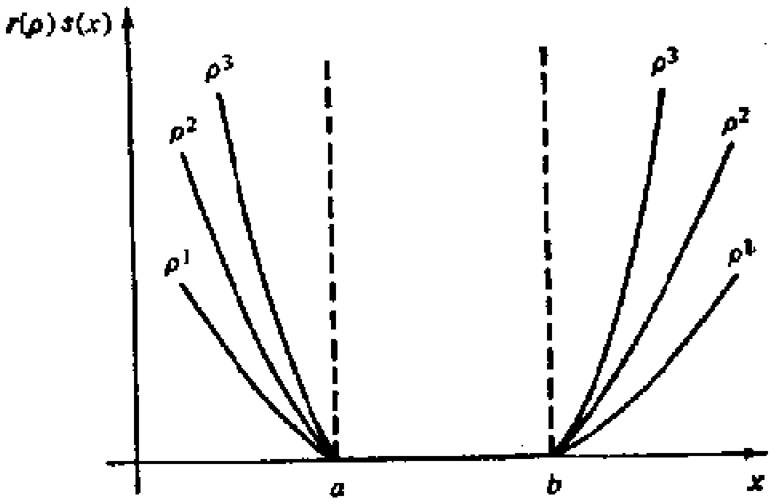


图 12.1 外部惩罚函数

在转向外部惩罚函数法的收敛性以前，先考虑一个小例题.

例 12.1.1

求$f(x)=(x)^2(x\in R)$的极小值点，受限制于约束 $x\geqslant 1$. 最优解显然是 $x^*=1$. 以 $\alpha=2$ 构成形如(12.14)的增广目标函数，这样就有无约束最优化问题

$$\min_{x\in R} F(x, \rho^k)=(x)^2+\frac{1}{\rho^k}[\min(0, x-1)]^2. \tag{12.18}$$

对任何给定的 $\rho^k>0$, 函数 F 是凸的，它的极小值点是

$$x^{k*}=\frac{1}{\rho^k+1}. \tag{12.19}$$

注意，对任何 $\rho^k>0$, 这个点关于原来的问题是非能行的. 随着$\{\rho^k\}\to 0$, 点 x^{k*} 从能行集外部趋于 x^*. 当然，在任何实际问题中，求 $F(x, \rho^k)$的无约束极小必须采用某个数值算法，如在前几章中提出的算法. ▌

现在叙述并证明外部惩罚函数法的某些收敛性结果. 首先需要

引理 12.1

设 F 由(12.17)给出，并令

$$\rho^k \geqslant \rho^{k+1} > 0. \tag{12.20}$$

假定 $F(x, \rho^k)$ 和 $F(x, \rho^{k+1})$ 分别在 x^{k*} 和 x^{k+1*} 达到它们在 R^n 上的极小值. 则

$$F(x^{k+1*}, \rho^{k+1}) \geqslant F(x^{k*}, \rho^k), \tag{12.21}$$

$$s(x^{k*}) \geqslant s(x^{k+1*}), \tag{12.22}$$

$$f(x^{k+1*}) \geqslant f(x^{k*}). \tag{12.23}$$

【证明】 因为 $r(\rho)s(x) \geqslant 0$, 且 r 随 ρ 的减少而增加，我们有

$$F(x^{k+1*}, \rho^{k+1}) = f(x^{k+1*}) + r(\rho^{k+1}) s(x^{k+1*}) \tag{12.24}$$

$$\geqslant f(x^{k+1*}) + r(\rho^k) s(x^{k+1*}) \tag{12.25}$$

$$\geqslant f(x^{k*}) + r(\rho^k) s(x^{k*}) = F(x^{k*}, \rho^k), \tag{12.26}$$

其中最后一个不等式由 x^{k*} 使 $F(x, \rho^k)$ 达到极小而推出. 因此(12.21)成立. 联合(12.25)和(12.26)，得出

$$f(x^{k*}) + r(\rho^k) s(x^{k*}) \leqslant f(x^{k+1*}) + r(\rho^k) s(x^{k+1*}). \tag{12.27}$$

由 x^{k+1*} 的定义，又有

$$f(x^{k+1*}) + r(\rho^{k+1}) s(x^{k+1*}) \leqslant f(x^{k*}) + r(\rho^{k+1}) s(x^{k*}). \tag{12.28}$$

把最后两个不等式相加，我们得到

$$r(\rho^k)[s(x^{k*}) - s(x^{k+1*})] \leqslant r(\rho^{k+1})[s(x^{k*}) - s(x^{k+1*})]. \tag{12.29}$$

再因 $r(\rho^k) \leqslant r(\rho^{k+1})$, 不等式(12.22)成立. 从(12.22)和(12.27)导出

$$f(x^{k+1*}) - f(x^{k*}) \geqslant r(\rho^k)[s(x^{k*}) - s(x^{k+1*})] \geqslant 0, \tag{12.30}$$

从而(12.23)成立. ∎

现在，我们能证明外部惩罚函数法的基本收敛定理.

定理 12.2

假设问题(P)的能行集 X 非空，并存在一个 $\varepsilon > 0$ 使得集合

$$X^\varepsilon = \{x : x \in R^n,\ g_i(x) \geqslant -\varepsilon, \quad i = 1, \cdots, m; \\ |h_j(x)| \leqslant \varepsilon, \quad j = 1, \cdots, p\} \tag{12.31}$$

是紧致的. 又假设对所有 k, $F(x, \rho^k)$ 在 R^n 上达到它们的无约束

极小. 若$\{\rho^k\}$是趋于零的严格递减正数列,则存在(EP^k)的最优解的一个收敛子列$\{x^{k_i*}\}$, 并且任何这样的收敛子列的极限都是(P)的最优解.

【证明】 由引理 12.1, $F(x^{k*}, \rho^k)$是一个递增序列. 因为X是紧致的且f是连续的, 所以至少存在一点$x^*\in X$, 使f在x^*达到它的极小值, 即x^*是(P)的最优解. 于是, 对$k=0, 1, \cdots$有

$$f(x^*)=f(x^*)+r(\rho^k)s(x^*)\geqslant F(x^{k*}, \rho^k), \tag{12.32}$$

从而序列$\{F(x^{k*}, \rho^k)\}$有上界. 因此, 它收敛于极限F^0(参见[2]). 类似地, 序列$\{f(x^{k*})\}$是递增的, 并且

$$f(x^{k*})\leqslant f(x^{k*})+r(\rho^k)s(x^{k*})=F(x^{k*}, \rho^k). \tag{12.33}$$

由(12.32)和(12.33), 我们得到

$$f(x^{k*})\leqslant f(x^*), \tag{12.34}$$

从而$\{f(x^{k*})\}$收敛于极限f^0. 而且,

$$\lim_{k\to\infty}\{r(\rho^k)s(x^{k*})\}=F^0-f^0. \tag{12.35}$$

由(12.16)推出

$$\lim_{k\to\infty}\{s(x^{k*})\}=0. \tag{12.36}$$

从(12.11), (12.12)和(12.36)我们断定: 对每个$\delta>0$, 存在一个自然数$K(\delta)$, 使得对$k\geqslant K(\delta)$有$x^{k*}\in X^\delta$. 于是, 对充分大的$\hat{K}(\varepsilon)$, 所有$k\geqslant\hat{K}(\varepsilon)$的点$x^{k*}$将落在紧致集$X^\varepsilon$中. 因此存在一个子序列$\{x^{k_i*}\}$收敛于极限$x^0$, 并推得$s(x^0)=0$. 这样, $x^0\in X$. 从x^*的最优性, 我们得到

$$f(x^*)\leqslant f(x^0). \tag{12.37}$$

现在, 对所有收敛子列中的k_i, 有

$$\begin{aligned}f(x^{k_i*})&\leqslant f(x^{k_i*})+r(\rho^{k_i})s(x^{k_i*})\\&\leqslant f(x^*)+r(\rho^{k_i})s(x^*)=f(x^*).\end{aligned} \tag{12.38}$$

于是

$$\lim_{k_i\to\infty}\{f(x^{k_i*})\}=f(x^0)\leqslant f(x^*). \tag{12.39}$$

由(12.37)和(12.39), 我们最终得到

$$f(x^0)=f(x^*), \tag{12.40}$$

并且 x^0 必定是(P)的一个解. ∎

Zangwill[52] 研究了使 $F(x, \rho^k)$ 达到它的极小值的条件，这是用于收敛性定理的一种假定. 他找到了两个这样的条件. 第一个说，f 是一个每当 $\|x\|\to+\infty$ 时必有 $f(x)\to+\infty$ 的函数; 另一个包含着同样推论的条件是，对某个 $\varepsilon>0$, X^ε 是紧致的，并且 $F(x, \rho)$ 是 x 的凸函数. 这个条件无疑可减弱到广义凸函数类(第6章). 注意，收敛性定理涉及到(P)的整体最优解，这样它似乎是很强的结果. 然而，那里假定了对每个 k，我们求得无约束规划 (EP^k) 的整体最优解，如同我们在前几章所看到，那是一个相当困难的任务. 基于这个理由，这个方法主要适用于凸规划(或它们的推广)，以及其增广目标函数至少是强拟凸的规划. 无约束规划 (EP^k) 中整体极小的存在性问题已由 Evans、Could[10] 研究过，他们导出了存在这种极值的充要条件. 很遗憾，这种条件的检验对大多数问题难以实现.

用外部惩罚函数算法收敛于非凸规划的局部极小，这结果也能在略有不同的假定下得到. 有兴趣的读者可参考 Fiacco、McCormick[15]，这里仅提及他们的主要结果. 记 X(能行集) 中的点集 A 是使目标函数 f 取极小值 v^* 的局部极小值点的集合，假设 A 是满足一个温和的正则性条件的非空紧致集，则存在一个紧致集 S，使 A 嵌在 S 的内部，而且，对充分大的 k，(EP^k) 的无约束极小值是在 S 的内部达到. 此外，

$$\lim_{k\to\infty}\{f(x^{k*})\}=v^*, \tag{12.41}$$

并且 $\{x^{k*}\}$ 的任何收敛子列的每个极限点都在 A 中.

最后，一些有效无约束数值方法的成功应用也要求连续可微性，这本身就对某些惩罚函数提出了问题，即使在(P)中出现的所有函数是连续可微的也如此. 这个问题在本章后面讨论. 后面要包含的另一方面内容是：对于满足强相容性或 Slater 条件(参看 4.5 节)的凸规划，单个无约束极小化能代替一系列极小化而获得所求的约束最优解.

12.2 内部惩罚函数

这里，不等式约束的非线性规划，是通过一系列无约束最优化问题来求解的，它们的极小值点严格地满足约束，即落在能行集的内部。我们将看到，当从内部穿越能行集的边界时，“障碍函数”加以无限大值的惩罚，这种障碍函数的建立，保证了所求得的无约束极小值点始终逗留在能行集的内部。因为算法要求能行集内部非空，所以等式约束问题不能用下面描述的方法处理。但是，另一些内部型惩罚函数法已经导出，它们能够求解带有等式约束的如上节的规划(P)那样的问题，这种方法将在后面提及。

所以，考虑非线性规划

(PI) $$\min f(x) \tag{12.42}$$

受限制于约束

$$g_i(x) \geqslant 0, \quad i=1, \cdots, m. \tag{12.43}$$

与前相同，假设 f, g_1, $\cdots$, g_m 是 R^n 上的连续函数。用 X 表示规划(PI)的能行集，即满足(12.43)的 $x \in R^n$ 的集合。令 X^0 是集 X 的内部。假设成立关于规划(PI)的下列正则性条件：

1. 集 X 是闭的，X^0 非空，且 X 是 X^0 的闭包；

2. 存在一个具有 $f(x^0)=\alpha^0$ 的点 $x^0 \in X$，使得集合 $S(f, \alpha^0) \cap X$ 是紧致的，其中 $S(f, \alpha^0)$是 f 在 α^0 的水平集。

在惩罚函数的表述中，令 q 是 $x \in R^n$ 的实值函数，并且 q 在 X^0 的每一点上连续。又若 $\{x^k\}$ 是 X^0 中收敛于 X 边界上某个 $\hat{x}$ 的任何点列，亦即

$$I(\hat{x}) = \{i : g_i(\hat{x}) = 0\} \neq \emptyset, \tag{12.44}$$

则

$$\lim_{k \to \infty} \{q(x^k)\} = +\infty. \tag{12.45}$$

再令 t 是 $\rho \in R$ 的实值函数，使得

$$\rho^1 > \rho^2 > 0 \quad 蕴涵 \quad t(\rho^1) > t(\rho^2) > 0, \tag{12.46}$$

$$\lim_{k \to \infty} \{\rho^k\} = 0 \quad 蕴涵 \quad \lim_{k \to \infty} \{t(\rho^k)\} = 0. \tag{12.47}$$

函数 $t(\rho^k)q(x)$ 称为**内部惩罚函数，或障碍函数.**

内部惩罚函数法可叙述如下：对于 $k=0, 1, \cdots$，定义 G 为在一系列无约束最优化问题 (IP^k) 中求极小的增广目标函数，(IP^k) 由下式给出：

$$(\mathrm{IP}^k) \qquad \min_x G(x, \rho^k)=f(x)+t(\rho^k)q(x). \tag{12.48}$$

令 $x^0\in X^0$ 是出发点，并对 ρ^0 指定一个正值．从 x^0 出发，用某个无约束极小化方法解问题(IP^0)，并记 x^{0*} 是(IP^0)的一个解．可预知 $x^{0*}\in X^0$．减小 ρ^0 到 ρ^1，从 x^{0*} 出发求解问题(IP^1)，记(IP^1)的最优解为 x^{1*}．用这种方式，对一个严格递减序列 ρ^k，总是从 x^{k-1*} 出发继续求解(IP^k)．在已叙述过的关于(PI)的假定下，这个方法能找到一个最优解，我们将在后面予以证明.

对于函数 t 和 q，最通常的选择是

$$t_1(\rho)=\rho, \tag{12.49}$$

$$t_2(\rho)=(\rho)^2, \tag{12.50}$$

$$q_1(x)=-\sum_{i=1}^{m}\log g_i(x), \tag{12.51}$$

$$q_2(x)=\sum_{i=1}^{m}\frac{1}{g_i(x)}, \tag{12.52}$$

$$q_2(x)=\sum_{i=1}^{m}\frac{1}{[g_i(x)]^2}, \tag{12.53}$$

$$q_4(x)=\sum_{i=1}^{m}\frac{1}{\max[0,\ g_i(x)]}. \tag{12.54}$$

例 12.2.1

考虑下列单变量的小型问题：

$$\min f(x)=\frac{1}{2}x \tag{12.55}$$

受限制于

$$g(x)=x-1\geqslant 0. \tag{12.56}$$

显然最优解在 $x^*=1$ 并且 $f(x^*)=\frac{1}{2}$．假设对于障碍函数，我们选择上面给出的 t_2 和 q_2，那么，

$$G(x, \rho^k)=\frac{1}{2}x+(\rho^k)^2(x-1)^{-1}. \tag{12.57}$$

读者容易验证 G 的无约束极小由下式给出:

$$x^{k*}=1+\sqrt{2}\rho^k, \tag{12.58}$$

$$f(x^{k*})=\frac{1}{2}+\frac{\rho^k}{\sqrt{2}}. \tag{12.59}$$

这样,最优的无约束极小值点全在 X^0 中, 而且随着 ρ^k 值的相继减小而收敛于 x^*. 原问题及几个 ρ^k 值的无约束问题示明在图12.2中. 】

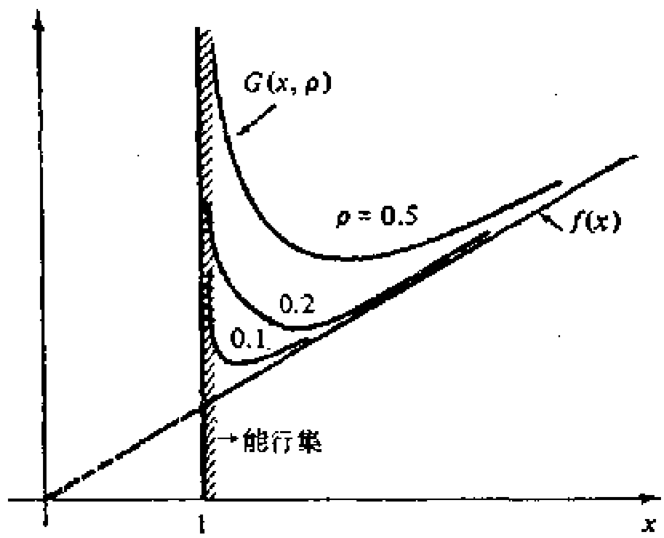


图 **12.2** 内部惩罚法

为了给出这个方法的一个简单的收敛性证明, 假定所用的函数 q 当 $x\in X^0$ 时为正. 这个假定对于对数惩罚 q_1 可能不成立, 然而也可证明利用这类惩罚的内部惩罚法的收敛性.

定理 12.3

假定上述的关于规划(PI)的正则性条件(1)和(2)满足, 而且在增广目标函数中用到的各个函数 q 当 $x\in X^0$ 时均为正. 假设对所有 k, $G(x, \rho^k)$在 X^0 中达到它的无约束极小值. 如果$\{\rho^k\}$是一收敛于零的严格递减正数列, 则存在(IP^k)的最优解的收敛子列$\{x^{k_j*}\}$, 并且任何这样的收敛子列的极限都是(PI)的最优解.

【证明】 由条件(2), 集合 $S(f, \alpha^0)\cap X$ 是非空的紧致的, 因此连续函数 f 在某个点 $x^*\in X$ 达到它在 X 上的极小值 $f(x^*)$. 令

$G(x^{k*}, \rho^k)$是规划(IP^k)中增广目标函数的极小值. 则由(12.46)和(12.48),

$$G(x^{0*}, \rho^0) \geqslant G(x^{1*}, \rho^1) \geqslant \cdots \geqslant f(x^*). \tag{12.60}$$

因为$\{G(x^{k*}, \rho^k)\}$是有下界的严格递减序列, 它收敛于极限$\hat{G} \geqslant f(x^*)$. 假设$\hat{G} > f(x^*)$. 由条件(1)和f的连续性,我们断定,存在一个数$\delta > 0$和球形邻域$N_\delta(x^*)$,使得$X^0 \cap N_\delta(x^*) \neq \emptyset$, 且对一切$x \in N_\delta(x^*)$有

$$f(x) \leqslant \hat{G} - \frac{1}{2}[\hat{G} - f(x^*)]. \tag{12.61}$$

在$X^0 \cap N_\delta(x^*)$中取任何点$\bar{x}$. 从(12.46)和(12.47)得出, 存在一个自然数K,使得对每个$k \geqslant K$,有

$$t(\rho^k) q(\bar{x}) < \frac{1}{4}[\hat{G} - f(x^*)]. \tag{12.62}$$

于是,对$k \geqslant K$有

$$G(x^{k*}, \rho^k) \leqslant f(\bar{x}) + t(\rho^k) q(\bar{x}) < \hat{G} - \frac{1}{4}[\hat{G} - f(x^*)], \tag{12.63}$$

这与$\{G(x^{k*}, \rho^k)\}$单调地收敛于$\hat{G}$矛盾. 因此,$\hat{G} = f(x^*)$.

由条件(2), 存在一个$\hat{K}$, 使得对所有$k \geqslant \hat{K}$, 点x^{k*}落在一个紧致集中. 从而存在一子列$\{x^{k_i*}\}$收敛于极限$\hat{x} \in X$. 假设$\hat{x}$不是(PI)的最优解,则$f(\hat{x}) > f(x^*)$, 以致序列$\{(f(x^{k_i*}) + t(\rho^{k_i}) q(x^{k_i*}) - f(x^*)\}$并不收敛于零,这样便与

$$\lim_{k \to \infty} \{G(x^{k*}, \rho^k)\} = f(x^*) \tag{12.64}$$

矛盾,所以必有$f(\hat{x}) = f(x^*)$,从而$\hat{x}$是(PI)的最优解. ∎

在这个定理中, 最重要的假定或许是$G(x, \rho^k)$在X^0中达到它们的极小值,或等价地说,问题(IP^k)在X^0中有最优解. 现在给出确保这些最优值存在的一个充分条件.

如果正则性条件(1)成立,并且X的内部X^0由下式给出:

$$X^0 = \{x : x \in R^n, \ g_i(x) > 0, \ i = 1, \cdots, m\} \neq \emptyset, \tag{12.65}$$

则规划(PI)定义为强相容的.

于是我们有下述引理.

引理 12.4

假定 $X\subset R^n$ 是紧致的，并且规划 (PI) 是强相容的，则 $G(x, \rho^k)$在 X^0 中达到它们的无约束极小值.

【证明】 令

$$\inf_{x\in X^0} G(x, \rho)=\alpha, \tag{12.66}$$

那末由下确界的定义，存在点 $x^i\in X^0$ 的一个序列$\{x^i\}$，使得

$$\lim_{i\to\infty}\{G(x^i, \rho)\}=\alpha. \tag{12.67}$$

因为$\{x^i\}$包含在紧致集 X 中，它有收敛子列$\{x^{i_j}\}$，从而

$$\lim_{i_j\to\infty}\{x^{i_j}\}=\hat{x}\in X. \tag{12.68}$$

假定 $\hat{x}\in X^0$. 那么，由 G 的连续性，因为$\{G(x^i, \rho)\}$的任何子列必收敛于 α，我们有

$$\lim_{i_j\to\infty}\{G(x^{i_j}, \rho)\}=\lim_{i_j\to\infty}\{f(x^{i_j})\}+\lim_{i_j\to\infty}\{t(\rho)q(x^{i_j})\} \tag{12.69}$$

$$=f(\hat{x})+t(\rho)q(\hat{x})=\alpha. \tag{12.70}$$

因此

$$G(\hat{x}, \rho)=\min_{x\in X^0} G(x, \rho). \tag{12.71}$$

现在假定 $\hat{x}\overline{\in} X^0$. 那么 $\hat{x}$ 是在 X 的边界上，由 (12.45) 和 (12.69)，并由于 $t(\rho)$为正，我们得到

$$\inf_{x\in X^0} G(x, \rho)=f(\hat{x})+\lim_{i_j\to\infty}\{t(\rho)q(x^{i_j})\}=+\infty, \tag{12.72}$$

这是一个矛盾，从而 $\hat{x}\in X^0$. ∎

用内部惩罚法收敛到非凸规划的局部极小值点，这个结果在较上面弱一点的假定下也能证明(参见 Fiacco、McCormick[15]).

内部惩罚法是以 Carroll[6] 提出的一个想法为基础的，他把一个约束非线性规划变换为一系列无约束极小化问题，他所用的函数 t_1 和 q_2 分别由(12.49)和(12.52)给出. Carroll 的想法随后被 Fiacco、McCormick[12, 15] 定型、推广，作了彻底的研究，他们发展了序列无约束极小化方法 (SUMT)，这或许是至今最著名的惩罚函数法.

在规划(PI)是标准凸规划并且 $f, g_1, \cdots, g_m$ 是连续可微的情况下,SUMT 方法具有某些有趣的原有-对偶特征. 考察(IP^k),即第 k 次无约束极小化问题:

$$(\text{IP}^k) \qquad \min_x G(x, \rho^k) = f(x) + \rho^k \sum_{i=1}^m \frac{1}{g_i(x)}. \tag{12.73}$$

若 $x^{k*} \in X^0$ 使(IP^k)达到极小,则

$$\nabla G(x^{k*}, \rho^k) = \nabla f(x^{k*}) - \rho^k \sum_{i=1}^m \frac{\nabla g_i(x^{k*})}{[g_i(x^{k*})]^2} = 0. \tag{12.74}$$

取(PI)的 Lagrange 式

$$L(x, \lambda) = f(x) - \sum_{i=1}^m \lambda_i g_i(x), \tag{12.75}$$

建议定义 Lagrange 乘子为

$$\lambda_i^k = \frac{\rho^k}{[g_i(x^{k*})]^2}, \quad i = 1, \cdots, m. \tag{12.76}$$

λ_i^k 的这种选择,使我们有

$$\nabla G(x^{k*}, \rho^k) = \nabla L(x^{k*}, \lambda^k) = 0. \tag{12.77}$$

注意,虽然 $\lambda_i^k \geqslant 0$,但因为

$$\lambda_i^k g_i(x^{k*}) = \frac{\rho^k}{g_i(x^{k*})} \neq 0, \quad i = 1, \cdots, m, \tag{12.78}$$

定理 4.41 中叙述的关于凸规划的 Kuhn-Tucker 必要条件,对于(PI)在点 x^{k*} 并不成立. 然而能够证明,向量 $\lambda^k = (\lambda_1^k, \cdots, \lambda_m^k)$ 关于(PI)的对偶规划是能行的,其中 λ_i^k 由(12.76)定义. 这个对偶规划(参见 5.4 节)由下式给出:

$$(\text{DPI}) \quad \max_{\lambda \geqslant 0} \left\{ \inf_x \left[f(x) - \sum_{i=1}^m \lambda_i g_i(x) \right] \right\} = \max_{\lambda \geqslant 0} \{ \inf_x L(x, \lambda^k) \}. \tag{12.79}$$

由 f 的凸性、g_i 的凹性及 λ_i^k 的定义和(12.77),可推出

$$L(x^{k*}, \lambda^k) = \inf_x L(x, \lambda^k), \tag{12.80}$$

并且 λ^k 对于(DPI)是能行的. 所以,由定理 5.5,值 $L(x^{k*}, \lambda^k)$ 是规划(PI)中 f 的极小值的一个下界. 随着 x^{k*} 趋于(PI)的最优解 x^*,$f(x^{k*})$ 与 $L(x^{k*}, \lambda^k)$ 之差越来越小直至二者达到相等. 这样,

(IP^k)的每一个成功的解，也提供了$f(x^{k*})$离开极小值有多远的一个估计．这个信息对于终止无约束极小序列可能是有用的．

正如已注意到的，等式约束非线性规划不能用内部惩罚法求解．然而，我们能利用**混合惩罚法**，去求解如(12.3)至(12.5)给出的等式-不等式约束问题．增广目标函数定义为

$$H(x,\ \rho,\ \eta)=f(x)+t(\rho)q(x)+r(\eta)s(x), \qquad (12.81)$$

其中函数t、q、r和s如上面对“纯粹”的外部惩罚法和内部惩罚法那样定义，不过$q(x)$和$s(x)$分别仅应用于不等式约束和等式约束．例如，我们可取

$$H(x,\ \rho,\ \eta)=f(x)-\rho\sum_{i=1}^{m}\log g_i(x)+\frac{1}{\eta}\sum_{j=1}^{m}(h_j(x))^2. \qquad (12.82)$$

混合惩罚法由解一系列无约束极小化问题组成：

$$(\mathrm{MP}^k)\quad \min_x H(x,\ \rho^k,\ \eta^k)=f(x)+t(\rho^k)q(x)+r(\eta^k)s(x). \qquad (12.83)$$

在适当的假定下，序列$\{H(x^{k*},\rho^k,\eta^k)\}$收敛于(P)的最优值$f(x^*)$，并且$\{x^{k*}\}$的子序列收敛于$x^*$，参见[13, 15]．

12.3 无参数惩罚法

实现惩罚函数法的一个尚未解决的问题涉及参数ρ的选择．必须决定参数的初始值ρ^0，并决定一个修改ρ值的规则，以便得到一个单调递减趋于零的序列．为避免选择参数时左右为难，可借助于修改惩罚函数来使参数自动选取，或等价地说，把方法修改成不需要参数的．下面就要看到，外部惩罚法和内部惩罚法二者都能作这样的修改．先从外部惩罚法开始．令x^*是目标函数f在规划(P)的由(12.6)给出的能行集X上的整体极小值点，假设我们有了关于f在点x^*的下界的一个估计值ω^0，即

$$\omega^0\leqslant f(x^*). \qquad (12.84)$$

又假设我们求解如下无约束最优化问题：

(EPF^0) $\min_{x\in R^n} \hat{F}(x, \omega^0)$

$$=\psi(\omega^0-f(x))+\sum_{i=1}^{m}\psi(g_i(x))+\sum_{j=1}^{p}\zeta(h_j(x)), \quad (12.85)$$

其中 ψ 和 ζ 分别由(12.7)和(12.8)定义. 令 x^{0*} 是(12.85)的最优解. 若(P)的最优解 x^* 恰好是 f 在 R^n 上的无约束极小值点, 则 $x^{0*}=x^*$, 从而算法终止. 否则, 只要类似于 12.1 节中叙述的正则性条件满足,则

$$\omega^0\leqslant f(x^{0*})\leqslant f(x^*), \quad (12.86)$$

并按下述方式进行.

令

(EPF^k) $\min_{x\in R^n} \hat{F}(x, \omega^k)$

$$=\psi(\omega^k-f(x))+\sum_{i=1}^{m}\psi(g_i(x))+\sum_{j=1}^{p}\zeta(h_j(x)), \quad (12.87)$$

其中$\{\omega^k\}$是严格递增序列, 它的元素 ω^k 由 ω^{k-1} 及(EPF^{k-1})的最优解算出. 对 $k=1, 2, \cdots$, 求解(12.87),如 12.1 节中那样, 我们得到一个点列 $\{x^{k*}\}$, 它包含收敛于(P)的最优解的子列. 进一步的细节见[29, 33, 36].

考虑由(12.42)和(12.43)给出的问题(PI), 以及对内部惩罚法之族叙述的正则性条件. 无参数的内部惩罚法建基于对 $k=1, 2, \cdots$, 求解如下一系列无约束最优化问题:

$$(\text{IPF}^k) \quad \min_{x\in R^n} \hat{G}(x, x^{k-1*})=\frac{1}{f(x^{k-1*})-f(x)}+\sum_{i=1}^{m}\frac{1}{g_i(x)}, \quad (12.88)$$

其中 x^{0*} 是在 X^0 中的一个任意点,x^{k-1*} 是(IPF^{k-1})的最优解. 这个方法称为**无参数 SUMT**[14]. 在这方法中,内部惩罚法 SUMT 已作了修改. 取 q_1 如(12.51), 我们能构造另一种无参数的内部惩罚法,那是基于求解一系列问题

$$\min_{x\in R^n} \hat{G}_1(x, x^{k-1*})=-\log[f(x^{k-1*})-f(x)]-\sum_{i=1}^{m}\log g_i(x), \quad (12.89)$$

它等价于求解

$$\max_{x} [f(x^{k-1*})-f(x)] \prod_{i=1}^{m} g_i(x), \tag{12.90}$$

其中 x^{k-1*} 定义同上. 这后一方法,称为**中心方法**,由 Huard[26] 建立. 再者,在类似于 12.2 节叙述的条件下,无参数内部惩罚法收敛于(PI)的最优解. 能够证明,无参数法实际上等价于带有特定参数调节的普通惩罚函数法. 例如,SUMT 和它的无参数形式之间的关系可以明白地叙述如下:若 x^{k*} 是(12.88)给出的(IPF^k)的解,只要(PI)的一些正则性条件得到满足[15],则 x^{k*} 也是由(12.73)给出的(IP^k)的解,其中取

$$\rho^k = [f(x^{k-1*})-f(x^{k*})]^2. \tag{12.91}$$

类似的关系也能对其他无参数方法导出. Lootsma[33] 研究了这些方法的收敛速率,发现它们较之以前讨论的外部和内部惩罚算法没有特别的好处. 事实上,无参数法的收敛一般很慢,其中大多数是线性收敛速率. 为了改进收敛速率,提出了几种修改,比如给予(12.87)右边第一项以一个很小的权,这项包含目标函数 f,参见[29,36].

例 12.3.1

我们用无参数 SUMT 去解例 12.2.1 中提出的问题. 我们有

$$\min f(x) = \frac{1}{2}x \tag{12.92}$$

受限制于

$$g(x) = x-1 \geqslant 0. \tag{12.93}$$

在第 k 次迭代中,我们求解无约束最优化问题

$$\min_{x \in R} \{[f(x^{k-1*})-f(x)]^{-1}+(x-1)^{-1}\}. \tag{12.94}$$

容易证明,最优解 x^{k*} 由下式给出:

$$x^{k*} = \frac{1+(x^{k-1*}/\sqrt{2})}{1+(1/\sqrt{2})}. \tag{12.95}$$

因 $x^*=1$, 我们得到

$$\frac{x^{k*}-x^*}{x^{k-1*}-x^*} = \frac{1}{1+\sqrt{2}} \tag{12.96}$$

从而收敛速率是线性的. 有趣的是注意到,对于例 12.2.1 中所用的内部惩罚法,有

$$\frac{x^{k*}-x^*}{x^{k-1*}-x^*}=\frac{\rho^k}{\rho^{k-1}}, \tag{12.97}$$

因此,采用适当选择的 $\{\rho^k\}$,收敛能加速. ∎

12.4 恰当惩罚函数

在至今所讨论的一切惩罚方法中, 最优解是通过求解一系列无约束最优化问题得到的. 自然要问: 约束最优化问题能否通过进行单个无约束最优化就解决. 确实, 给定了一个约束非线性规划后, 若能找到一个实函数, 使它的无约束极小就是约束问题的解,那是很有益处的. 其结果是,在一定的场合寻找这样的函数是可能的,并且它们与本章已讨论的惩罚函数有关.

给定了一个非线性规划,例如 12.1 节中的规划(P),假设存在一个实函数,具有这样的性质,即它的单个无约束极小产生(P)的最优解. 因为这种函数常常属于惩罚函数的某些特定类型, 它们带有适当选择的参数,所以称之为**恰当惩罚函数**.

处理恰当惩罚函数的最早著作之一是 Zangwill[52]. 他指出: 对于一个满足强相容条件的凸规划,存在一个恰当惩罚函数,取包含一个充分小参数 ρ 的外部惩罚函数的形式, 这里我们在第 6 章广义凸性的观点下来推广 Zangwill 的结果.

考虑下列广义凸规划,其中使用了与第 6 章相适应的记号(对于不熟悉在那里为(h, ϕ)-凸函数引进的特殊记号的读者, 可把它们理解为通常的代数运算,然后把这些结果用于普通凸规划). 令

(GCP) $$\min f(x) \tag{12.98}$$

受限制于

$$g_i(x)\geqslant 0_\phi, \quad i=1, \cdots, m, \tag{12.99}$$

其中 f 是(h, ϕ)-凸函数, g_i 是(h, ϕ)-凹函数,都定义在一个弧式连通集 $S\subset R^n$ 上. 如前一样,令 $X\subset S$ 是满足(12.99)的点 $x\in R^n$ 的集合,并记 $X^0\subset X$ 是集合

$$X^0=\{x,\ x\in R^n,\ g_i(x)>0_\phi,\ i=1,\ \cdots,\ m\}. \qquad (12.100)$$

假定 X^0 非空，并设存在一个 $x^*\in X$ 是(GCP)的最优解. 定义

$$P(x,\ \rho)=f(x)[-]\frac{1}{\rho}[\cdot]\left\{\left[\sum_{i=1}^{m}\right]\min(0_\phi,\ g_i(x))\right\}, \quad (12.101)$$

其中

$$a[-]b-a[+]((-1)[\cdot]b)=\phi^{-1}[\phi(a)-\phi(b)], \quad (12.102)$$

$$\left[\sum_{i=1}^{m}\right]a_i=a_1[+]a_2[+]\cdots[+]a_m. \qquad (12.103)$$

若 f 是凸的，g_i 是凹的，我们得到

$$P(x,\ \rho)=f(x)-\frac{1}{\rho}\sum_{i=1}^{m}\min(0,\ g_i(x)). \qquad (12.104)$$

注意，$P(x,\ \rho)$的这个形式等价于 $\alpha=1$ 时由(12.7)定义的 ψ.

现在，我们叙述并证明下面的定理.

定理 12.5

存在一个 $\rho^*>0$，使得对所有 $\rho^*\geqslant\rho>0$，$P(x,\ \rho)$的无约束极小值点与(GCP)的最优解 x^* 重合.

【证明】 令 $\hat{x}\in X^0$，定义

$$\alpha=\phi(\min_i(g_i(\hat{x})))>0, \qquad (12.105)$$

$$\beta=\phi f(\hat{x})-\phi f(x^*)\geqslant 0. \qquad (12.106)$$

令

$$\frac{1}{\rho^*}=\frac{\beta+\varepsilon}{\alpha}>0, \qquad (12.107)$$

其中 $\varepsilon>0$ 是任意的.

其次，取任意点 $w\in S\subset R^n$ 使 $w\bar{\in}X$. 我们将证明，对每个 w 存在点 $\bar{x}\in X$，满足

$$P(\bar{x},\ \rho^*)<P(w,\ \rho^*). \qquad (12.108)$$

因为 $\bar{x}$ 是能行的，$f(\bar{x})=P(\bar{x},\ \rho^*)$. 而且，既然按假设(GCP)的最优解能达到，$P(x,\ \rho^*)$也必在(GCP)的一个最优点 $x^0\in X$ 处达到极小值. 令 $\bar{x}$ 是连结 $\hat{x}$ 和 w 的弧 $h_{\hat{x},w}$ 上的一点，即 $\bar{x}=h_{\hat{x},w}(\bar{\theta})$，$0<\bar{\theta}<1$，使得

$$I(\bar{x})=\{i:g_i(\bar{x})=0_\phi,\ i=1,\ \cdots,\ m\}\neq\emptyset. \tag{12.109}$$

换言之，$\bar{x}$ 在 X 的边界上．定义一个辅助函数

$$\tau(x)=f(x)[-]\frac{1}{\rho^*}[\cdot]\{[\sum_{i\in I(\bar{x})}]g_i(x)\}. \tag{12.110}$$

显然，$\tau(\bar{x})=P(\bar{x},\ \rho^*)=f(\bar{x})$，并且

$$\begin{aligned}[\sum_{i\in I(\bar{x})}]g_i(w)&=[\sum_{i\in I(\bar{x})}]\min(0_\phi,\ g_i(w))\\ &\geqslant\left[\sum_{i=1}^{m}\right]\min(0_\phi,\ g_i(w)).\end{aligned} \tag{12.111}$$

因此，

$$\tau(w)=f(w)[-]\frac{1}{\rho^*}[\cdot]\{[\sum_{i\in I(\bar{x})}]g_i(w)\} \tag{12.112}$$

$$\leqslant f(w)[-]\frac{1}{\rho^*}[\cdot]\left\{\left[\sum_{i=1}^{m}\right]\min(0_\phi,\ g_i(w))\right\}=P(w,\ \rho^*). \tag{12.113}$$

我们通过指出 $\tau(\bar{x})<\tau(w)$ 来完成证明．首先，我们有

$$\tau(\hat{x})=f(\hat{x})[-]\frac{\beta+\varepsilon}{\alpha}[\cdot]\{[\sum_{i\in I(\bar{x})}]g_i(\hat{x})\} \tag{12.114}$$

$$\leqslant f(\hat{x})[-]\frac{\beta+\varepsilon}{\alpha}[\cdot]\{\min_i g_i(\hat{x})\} \tag{12.115}$$

$$=f(\hat{x})[-]\phi^{-1}\left\{\frac{\beta+\varepsilon}{\alpha}\phi(\min_i g_i(\hat{x}))\right\} \tag{12.116}$$

$$=f(\hat{x})[-]\phi^{-1}(\beta+\varepsilon) \tag{12.117}$$

$$=\phi^{-1}(\phi f(x^*)-\varepsilon)<f(\bar{x})=\tau(\bar{x}). \tag{12.118}$$

其次，因为 τ 是 (h,ϕ)-凸函数，且对某个 $0<\bar{\theta}<1$ 成立 $\bar{x}=h_{\hat{x},w}(\bar{\theta})$，故推出

$$\tau(\bar{x})\leqslant((1-\bar{\theta})[\cdot]\tau(\hat{x}))[+](\bar{\theta}[\cdot]\tau(w)) \tag{12.119}$$

$$<((1-\bar{\theta})[\cdot]\tau(\bar{x}))[+](\bar{\theta}[\cdot]\tau(w)), \tag{12.120}$$

从而

$$\tau(\bar{x})=P(\bar{x},\ \rho^*)<\tau(w)\leqslant P(w,\ \rho^*), \tag{12.121}$$

因此(12.108)成立． ∎

虽然临界参数值 ρ^* 的存在性已在定理 12.5 中证实，但在实际上，因为 ρ^* 依赖于 $f(x^*)$，所以只有在问题本身解出后 ρ^* 才能

计算．然而，倘若 $f(x^*)$ 的下界能找到，如在许多具体情况下那样，我们就能够估计 ρ^*．令 f^c 是这样一个下界，定义

$$\beta^e = \phi f(\hat{x}) - \phi f^e \geqslant \beta, \tag{12.122}$$

$$\frac{1}{\rho^e} = \frac{\beta^e + \varepsilon}{\alpha} \geqslant \frac{1}{\rho^*}, \tag{12.123}$$

则 ρ^e 能用来代替恰当惩罚函数中的 ρ^*．

例 **12.4.1**

下面的简单例子说明了恰当惩罚函数的思想：

$$\min f(x) = x^{1/2} \tag{12.124}$$

受限制于

$$g(x) = \frac{1}{4} x \geqslant 1. \tag{12.125}$$

从第 6 章和第 7 章得知，f 和 g 分别是 (log, log)-凸函数和(log, log)-凹函数．换句话说，我们有一个通常的几何规划问题．

能行集和目标函数在图 12.3 中示出．显然，最优解是在 $x^* = 4$，我们通过寻找临界参数 ρ^*（利用对 x^* 准确位置的知识）着手构造恰当惩罚函数．取 $\hat{x} = 4e$，这个点是能行的．由(12.105)至(12.107)，并令 $\varepsilon = \dfrac{1}{2}$，得到

$$\alpha = \log e = 1, \tag{12.126}$$

$$\beta = \log(4e)^{1/2} - \log(4)^{1/2} = \frac{1}{2}, \tag{12.127}$$

$$\frac{1}{\rho^*} = 1. \tag{12.128}$$

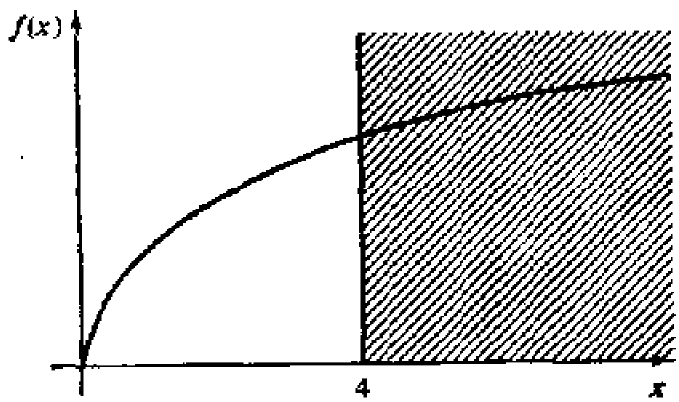


图 **12.3**　约束最优化问题

因此由(12.101),

$$P(x,\ \rho^*)=(x)^{1/2}[-]\frac{1}{\rho^*}[\cdot]\min\left(1,\frac{x}{4}\right) \tag{12.129}$$

$$=\exp\left\{\frac{1}{2}\log x-\min\left(0,\log\frac{x}{4}\right)\right\}, \tag{12.130}$$

从而,

$$P(x,\ \rho^*)=\begin{cases}\dfrac{4}{(x)^{1/2}}, & x\leqslant 4,\\ (x)^{1/2}, & x\geqslant 4.\end{cases} \tag{12.131}$$

这个函数示明在图 12.4 中. 这个惩罚函数的无约束极小值点是 $x^*=4$, 所以是一个恰当惩罚函数. 注意, $P(x,\ \rho^*)$ 在 x^* 不可微, 这个计算上不希望有的特性, 是某些恰当惩罚函数构造中所固有的. ▮

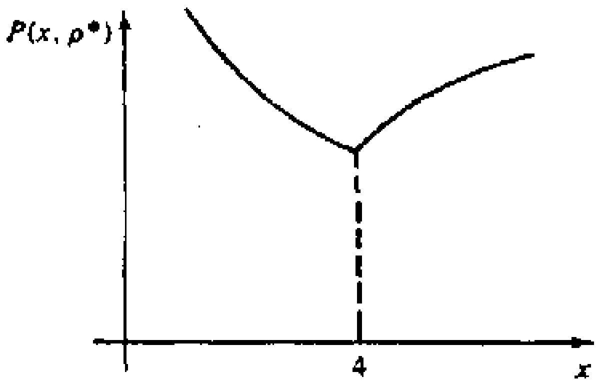


图 12.4 恰当惩罚函数

Pietrzykowski[46] 利用 (12.104) 给出的恰当惩罚函数 $P(x,\ \rho)$, 求得了把非凸规划的约束局部极小与 $P(x,\ \rho)$ 的无约束极小联系起来的一些结果. Zangwill-Pietrzykowski 恰当惩罚函数法的主要缺陷之一, 如同刚才例中所见的, 是 $P(x,\ \rho)$ 在一个使 $I(\hat{x})\neq\emptyset$——即对某个 i 成立 $g_i(\hat{x})=0$——的点 $\hat{x}\in R^n$ 上不可微, 即使 f 和 g_i 是连续可微也一样. 因此, 象第10章和第11章中那些有效的无约束极小化方法就不能直接应用. 对于这类分片可微惩罚函数, 专门的最优化方法已由 Conn[7] 和 Conn-Pietrzykowski[8] 提出.

在 Evans、Could、Tolle[11] 的新近工作中, 对一大类分片可

微函数，建立了一个十分一般的理论．他们的理论也可以看作是联系惩罚函数与 Lagrange 式的环节(这个关系将在下一节更详细地讨论)．对应于每个不等式约束 $g_i(x)\geqslant 0$，Evans、Gould 和 Tolle 定义数 β_i 和向量 $\lambda\in A$ 的一个**乘子函数** γ_i，其中 A 是 R^k 的子集，使得 γ_i 的值域是 $R\cup\{-\infty\}$．还假定，对每个固定的 $\lambda\in R^k$，γ_i 是 β_i 的连续非减函数．然后，定义**扩充 Lagrange 式**为

$$P(x,\ \lambda)=f(x)-\sum_{i=1}^{m}\gamma_i(g_i(x),\ \lambda), \tag{12.132}$$

它联系着不等式约束的非线性规划，例如(12.42)和(12.43)给出的规划(PI)．与(PI)相联系的通常的 Lagrange 式就成为一个特殊情形，此时取 $k=m$、$A=\{\lambda:\lambda\in R^m,\ \lambda>0\}$且

$$\gamma_i(\beta_i,\ \lambda)=\lambda_i\beta_i,\quad i=1,\ \cdots,\ m. \tag{12.133}$$

于是，

$$P(x,\ \lambda)=f(x)-\sum_{i=1}^{m}\lambda_i g_i(x), \tag{12.134}$$

它就是在第 3 章中引进的 Lagrange 式．实际上，对于第 4 章定义的满足强相容条件(Slater 条件)的凸规划(CP)，Lagrange 式 $L(x,\ \lambda^*)$是一个恰当惩罚函数．那就是说，若“最优”Lagrange 乘子向量 λ^* 是已知的，则如同定理 **4.41** 所证明的，$L(x,\ \lambda^*)$的无约束极小值点与(CP)的约束最优解一致．(12.132)的另一个特殊情况为，取 $k=1$、$A=\{\lambda:\lambda\in R,\ \lambda>0\}$和

$$\gamma_i(\beta_i,\ \lambda)=\begin{cases}-\dfrac{\lambda_i}{\beta_i}, & \beta_i>0,\\ -\infty, & \beta_i\leqslant 0,\end{cases}\qquad i=1,\ \cdots,\ m. \tag{12.135}$$

因此，

$$P(x,\ \lambda)=\begin{cases}f(x)+\lambda\sum\limits_{i=1}^{m}\dfrac{1}{g_i(x)}, & x\in X,\\ +\infty, & x\overline{\in} X,\end{cases} \tag{12.136}$$

其中 X 与前面一样是(PI)的能行集．出现在(12.136)中的函数是 SUMT 方法中的内部惩罚函数，并且是由 Gould[23] 研究的“内部

"外部"类惩罚函数的一员. 最后, 令 $k=1$ 和 $A=\{\lambda:\lambda\in R, \lambda>0\}$, 并定义

$$\gamma_i(\beta_i, \lambda)=\lambda \min(0, \beta_i), \quad i=1, \cdots, m, \tag{12.137}$$

我们得到

$$P(x, \lambda)=f(x)-\lambda \sum_{i=1}^{m} \min[0, g_i(x)], \tag{12.138}$$

它就是上面谈及的关于凸规划的 Zangwill 恰当惩罚函数.

在以下的讨论中, 假定对每个 $\lambda\in A$ 有

$$\gamma_i(0, \lambda)=0, \quad i=1, \cdots, m. \tag{12.139}$$

注意, 这个条件排除(12.135)作为乘子函数的特殊情形. 现在我们能够用扩充 Lagrange 式来定义恰当惩罚函数, 我们说, $P(x, \lambda)$ 是下述规划的一个恰当惩罚函数:

(PI) $$\min f(x) \tag{12.140}$$

受限制于

$$g_i(x)\geqslant 0, \quad i=1, \cdots, m, \tag{12.141}$$

如果存在 $\lambda^*\in A$, 使得

$$\min_{x\in R^n} P(x, \lambda^*)=f(x^*), \tag{12.142}$$

其中 x^* 是 (PI) 的最优解. 因为 γ_i 对每个固定的 λ 是非减的, 从(12.139)得出, 对每个满足(12.141)的 $x\in R^n$, 成立

$$\sum_{i=1}^{m} \gamma_i(g_i(x), \lambda^*)\geqslant 0. \tag{12.143}$$

假设 x^* 是(PI)的解, 则

$$P(x^*, \lambda^*)=f(x^*)-\sum_{i=1}^{m} \gamma_i(g_i(x^*), \lambda^*)\leqslant f(x^*)=\min_{x\in R^n} P(x, \lambda^*), \tag{12.144}$$

这就是, x^* 是 $P(x, \lambda^*)$ 的一个无约束极小值点. 反之, 令 $\hat{x}\in R^n$ 满足(12.141), 使得对某个 $\lambda^*\in A$ 成立

$$f(\hat{x})=P(\hat{x}, \lambda^*)=\min_{x\in R^n} P(x, \lambda^*), \tag{12.145}$$

则对每个能行点 x,

$$f(\hat{x})\leqslant P(x, \lambda^*)\leqslant f(x). \tag{12.146}$$

换句话说，$\hat{x}$ 是(PI)的最优解.

Evans、Gould 和 Tolle[11] 曾研究某些特殊的乘子函数，对于它成立下列较强的结果：在关于 f 和 g_i 的某种温和的假定下，规划(PI)的最优解集合，和 λ 充分大时相应的恰当惩罚函数 $P(x,\lambda)$ 的整体极小解集合相同. 例如，一个这样的乘子函数是

$$\gamma_i(\beta_i,\lambda)=\min(0,\ e^{\lambda\beta_i}-1),\quad i=1,\ \cdots,\ m,\qquad(12.147)$$

其中 $\lambda\in\Lambda=\{\lambda:\lambda\in R,\ \lambda>0\}$，从而相应的扩充 Lagrange 式为

$$P(x,\ \lambda)=f(x)-\sum_{i=1}^{m}\min(0,\ e^{\lambda g_i(x)}-1).\qquad(12.148)$$

[11] 中最重要的结果，是用相应的扩充 Lagrange 式的无约束极小刻划 (PI) 的最优解的特征定理. 因为这些 Lagrange 式只是分片可微的，这种类型的恰当惩罚函数方法的有用性仍值得怀疑. 到现在为止，我们还没有求不可微函数极小值的行之有效的数值方法. 为此，我们离开这主题，以讨论 Fletcher 的恰当惩罚函数法作为结束，其中的惩罚函数对于有效地应用梯度型无约束极小化算法而言，是足够光滑的.

在一系列的著作中，Fetcher[16,17] 及 Fletcher-Lill[19] 首次导出了一种优美的方法，来对只含等式约束的非线性规划定义可微恰当惩罚函数. 这个方法后来又为 Fletcher 扩充到不等式约束问题[18]. 我们从等式约束情况着手来讲述 Fletcher 的方法. 考虑规划

$$\text{(PE)}\qquad \min f(x)\qquad(12.149)$$

受限制于

$$h_j(x)=0,\quad j=1,\ \cdots,\ p,\qquad(12.150)$$

其中 $f,\ h_1,\ \cdots,\ h_p$ 是 $x\in R^n$ 上二次连续可微函数且 $p\leqslant n$. 令 $J_p=\{1,\ \cdots,\ p\}$，对某个包含 q 个元素的子集 $J_q\subset J_p$，我们用下式规定在每个 $x\in R^n$ 上的 $n\times q$ 阵 N_q:

$$N_q(x)=[\nabla h_j(x),\ j\in J_q].\qquad(12.151)$$

N_q 的广义逆[44] 倘若存在，它定义为

$$N_q^+(x)=(N_q^TN_q)^{-1}N_q^T,\qquad(12.152)$$

为了表达简洁，在记号中对 x 的明显依赖省略了，又定义投影矩阵 $\hat{P}_q$ 为

$$\hat{P}_q = N_q N_q^+ = (N_q^+)^T N_q^T, \tag{12.153}$$

它把向量投影到由 ∇h_j, $j\in J_q$, 张成的子空间中，并且相补投影矩阵 P_q 为

$$P_q = I - \hat{P}_q, \tag{12.154}$$

其中 I 是单位阵. 矩阵 P_q 把向量投影到相切于由 q 个约束的交形成的流形上. 容易理解 P_q 和 $\hat{P}_q$ 也是 x 的函数. 关于投影矩阵的进一步细节在 13.2 节中给出.

Fletcher 给规划(PE)找到了整个一类恰当惩罚函数，其代表性的一个由下式给出：

$$\phi(x) = f(x) - (h(x))^T N_p^+ \nabla f(x) + \rho (h(x))^T N_p^+ (N_p^+)^T h(x), \tag{12.155}$$

其中 $h(x) = (h_1(x), \cdots, h_p(x))^T$. 我们来校核 ϕ 在(PE)的约束局部极小值点处的性态. 要求读者证明，采用上述记号，x^l 是局部极小值点的必要条件为

$$P_q \nabla f(x^l) = 0, \tag{12.156}$$

$$h(x^l) = 0, \tag{12.157}$$

$$\nabla \phi(x^l) = 0. \tag{12.158}$$

即，x^l 是 ϕ 的一个逗留点. 但是，从第 2 章我们知道，与(PE)联系的 Lagrange 式也必须以 x^l 为一个逗留点，即

$$\nabla L(x^l, \mu) = \nabla f(x^l) - N_p \mu = 0, \tag{12.159}$$

所以由(12.158)和(12.159)，Lagrange 乘子向量能写为

$$\mu = N_p^+ \nabla f(x^l). \tag{12.160}$$

我们可由 Lagrange 式来导出 ϕ. 令 μ 如同(12.160)中对每个 x 定义，我们得到一个函数

$$\psi(x) = f(x) - (h(x))^T N_p^+ \nabla f(x). \tag{12.161}$$

这个函数在(PE)的每个局部最优解处是逗留的，但是它不一定在异于 x^l 的点上与普通的 Lagrange 式相等（这 Lagrange 式中的 μ 与 x 无关）. 比较 ϕ 和 ψ，我们看出它们之间的差是在 ϕ 的最后一

项，这一项的作用是确保对充分大的 ρ，函数 ϕ 在(PE)的解上有一个极小值，而不仅是一个逗留点．事实上，Fletcher[16] 证明了下列结果．

定理 12.6

令 $x^* \in R^n$ 是这样一个点，它满足定理 2.8 中给出的、(PE)的严格局部极小值点的充分条件，则 $\nabla\phi(x^*)=0$，而且存在一个 $\rho^*>0$，使得对每个 $\rho \geqslant \rho^*$，矩阵 $\nabla^2\phi(x^*)$ 是正定的．

在我们能应用这个恰当惩罚函数法之前，还留下的唯一问题是要对 ρ 指定一个值，使单个无约束极小化会产生所求的(PE)的最优解．对于非线性约束函数，临界参数 ρ^* 的计算决不简单．Fletcher 和 Lill[19] 提议取

$$\rho=10\|\nabla^2 f(x^0)\|+r, \tag{12.162}$$

其中 x^0 是无约束极小化的出发点，并且初次取 $r=1$．然而，ρ 的这个值仅仅是 ρ^* 的一个近似，在极小化过程中，当一个负定的 $\nabla^2\phi(x)$ 出现时，就必须用增加 r 来调节 ρ．注意，若 $\nabla^2 f$ 不合用，就利用这个 Hesse 阵的一个有限差分近似．一个任意的、不一定很大的 ρ 值在极小化算法中可能造成数值计算的困难，所以应该避免．

由(12.155)给出的 ϕ，是在普通 Lagrange 式或其近似式上增添附加项而得到的，这样做是为了保证 ϕ 的 Hesse 阵在约束最优解处是正定的．象 ϕ 这样的函数也称为增广 Lagrange 式，将在下节更广泛地研究．

联系 ϕ 与 Lagrange 型函数的另一种方式是

$$\phi(x)=f(x)-(\gamma(x))^T h(x), \tag{12.163}$$

其中 $\gamma(x)\in R^p$ 由下式给出：

$$\gamma(x)=N_p^+\nabla f(x)-\rho N_p^+(N_p^+)^T h(x); \tag{12.164}$$

这里 $\gamma(x)$ 是 x 的函数，而在普通的 Lagrange 式中，向量 μ 与 x 无关．下面导出一个有趣且有用的结果．对任何给定的 $\hat{x}\in R^n$，考虑如下的二次规划：

(QP) $$\min Q(z)=z^T\nabla f(\hat{x})+\frac{1}{2}\rho z^Tz \qquad (12.165)$$

受限制于

$$l(z)=N_p^Tz+h(\hat{x})=0. \qquad (12.166)$$

于是，我们有下列定理。

定理 12.7

设(QP)所联系的 Lagrange 式定义为

$$L(z,\ \mu_Q)=Q(z)-\mu_Q^Tl(z), \qquad (12.167)$$

若 N_p 满秩，则

$$\mu_Q=\gamma(\hat{x}), \qquad (12.168)$$

其中 $\gamma(\hat{x})$ 由(12.164)给出。

这个定理的证明是线代数的一个练习，留给读者去做。注意，二次规划由出现在 (PE) 中的一些函数来定义，这些函数在 $\hat{x}$ 求值，所以在(12.166)中的 $l(z)$ 是 h 在 $\hat{x}$ 的一个线性近似，并且 $Q(z)$ 是 f 在 $\hat{x}$ 的一种"二次近似"，其中 $\nabla^2 f(\hat{x})$ 用 ρI 代替。

我们现在将看到，这样的二次规划的求解，是不等式约束规划的 Fletcher 恰当惩罚函数法的一个实质部分。

所以，考虑由(12.3)至(12.5)定义的问题(P)，并定义对应的二次规划

(QP1) $$\min Q(z)=z^T\nabla f(\hat{x})+\frac{1}{2}\rho z^Tz \qquad (12.169)$$

受限制于

$$l_g(z)=(N_m^g)^Tz+g(\hat{x})\geqslant 0, \qquad (12.170)$$

$$l_h(z)=(N_p^h)^Tz+h(\hat{x})=0, \qquad (12.171)$$

其中 $g(x)=(g_1(x),\ \cdots,\ g_m(x))^T$，$h(x)=(h_1(x),\ \cdots,\ h_p(x))^T$，并且 N_m^g 和 N_p^h 分别是以 $\nabla g_i(\hat{x})$ 和 $\nabla h_j(\hat{x})$ 为列的矩阵。一旦 ρ 值被指定并且出发点 $x^0\in R^n$ 被选出，(QP1)就能用某些二次规划算法来求解，它也得出 Lagrange 乘子向量 $\lambda_Q(\hat{x})$，$\mu_Q(\hat{x})$。假定在(QP1)的解处的积极不等式约束在(P)的解处也是积极的，还假定(P)和(QP1)的 Lagrange 乘子相同。现在，我们有惩罚函数

$$\phi(x)=f(x)-(\lambda_Q(\hat{x}))^T g(x)-(\mu_Q(\hat{x}))^T h(x), \quad (12.172)$$

它可借助于第 10 章或第 11 章中利用一阶导数或可能也用二阶导数的方法之一来求极小. 注意, (12.172) 的每次函数值计算需要求解二次规划(QP1). 然而, 由于它的结构, 二次规划的求解并不困难; 并且 Fletcher 的有限的试验[18] 表明, 一旦确定了(P)的局部极小值点上的积极不等式约束的正确集合, 就可以极少量的迭代次数达到收敛. 求解二次规划的算法将在后面几章中讨论.

12.5 乘子法和 Lagrange 法

实现头两节中叙述的方法的困难之一在本质上是固有的, 随着参数 ρ^k 或 $g_i(x^{k*})$ 的趋于零, 由于在能行集边界上的不连续性以及惩罚函数的 Hesse 阵变为病态, 数值计值更加困难. 因此, 导出那种只需要参数采取适度值的方法可能是有用的. 这里提出的方法就是基于这个事实, 但是我们在下面即可看到, 它们也与恰当惩罚函数和 Lagrange 式紧密联系.

首先考虑带有等式约束的非线性规划

(PE) $$\min f(x) \quad (12.173)$$

受限制于

$$h_j(x)=0, \quad j=1, \cdots, p. \quad (12.174)$$

为了求解这个规划, Hestenes[25] 提出下列有趣的惩罚函数法.

假设 $f, h_1, \cdots, h_p$ 是二次连续可微函数, 并设 $x^*\in R^n$ 是(PE)的一个局部解, 使得定理 2.8 给出的、严格局部极小值点的充分条件得以满足. 也就是说, 存在乘子 $\mu_1^*, \cdots, \mu_p^*$, 使得

$$\nabla f(x^*)-\sum_{j=1}^{p}\mu_j^*\nabla h_j(x^*)=0, \quad (12.175)$$

$$h_j(x^*)=0, \quad j=1, \cdots, p, \quad (12.176)$$

并且, 对每个满足下式的 $z\neq0$,

$$z^T\nabla h_j(x^*)=0, \quad j=1, \cdots, p, \quad (12.177)$$

我们有

$$z^T\nabla_x^2 L(x^*, \mu^*)z>0. \quad (12.178)$$

我们来证明，可能找到一个正数c^*，使x^*是增广 Lagrange 式$M(x, \mu^*)$的无约束极小值点，其中

$$M(x, \mu^*) = f(x) - \sum_{j=1}^{p} \mu_j^* h_j(x) + \frac{1}{2} c^* \sum_{j=1}^{p} (h_j(x))^2 \quad (12.179)$$

$$= L(x, \mu^*) + \frac{1}{2} c^* \sum_{j=1}^{p} (h_j(x))^2. \quad (12.180)$$

我们从下列结果开始，它属于 Arrow、Could、Howe[1].

引理 12.8

令u和v是紧致集$K \subset R^n$上的连续实函数，对所有$z \in K$，有$v(z) \geqslant 0$. 那末，每当$v(z) \leqslant 0$时$u(z) > 0$的充要条件是：存在一个数c^*，使得对所有$c \geqslant c^*$和$z \in K$成立

$$u(z) + cv(z) > 0. \quad (12.181)$$

【证明】 假设 (12.181) 对所有$z \in K$和$c \geqslant c^*$成立. 则$v(z) \leqslant 0$显然蕴涵$u(z) > 0$. 反之，令K'是使$u(z) \leqslant 0$的K的子集. 倘若$K' = \emptyset$，我们便完成了证明；于是假定K'非空，则

$$u(z) + cv(z) \geqslant u(z) > 0 \quad (12.182)$$

对所有$c \geqslant 0$及适合$z \overline{\in} K'$的$z \in K$成立. 因为K'是紧致的，函数u和v在K'上达到各自的极小值. 相应地用u^*和v^*记这些函数在K'上的极小值. 那末对某个$\hat{z} \in K'$，有

$$v^* = v(\hat{z}) \geqslant 0. \quad (12.183)$$

若$v(\hat{z}) = 0$，则由假设有$u(\hat{z}) > 0$，就与K'的定义矛盾. 因此

$$v^* = v(\hat{z}) > 0. \quad (12.184)$$

从而对一切$c \geqslant c^* > -u^*/v^*$和$z \in K'$，我们有

$$u(z) + cv(z) \geqslant u^* + cv^* > 0. \quad (12.185)$$

∎

推论 12.9

令A是一个$n \times n$阵，B是一个$m \times n$阵. 则对每个满足$Bz = 0$的$z \neq 0$都成立$z^T A z > 0$的充要条件是：存在一个数$c^* > 0$，使得对所有$c \geqslant c^*$和$z \neq 0$，成立

$$z^T (A + cB^T B) z > 0. \quad (12.186)$$

【证明】 在前面引理中令 $u(z)=z^TAz$ 和 $v(z)=(Bz)^TBz$，并令

$$K=\{z:z\in R^n,\ z^Tz=1\},\tag{12.187}$$

则这推论由引理 12.8 对 $z^Tz=1$ 的所有 z 成立而推得，从而它对所有 $z\neq0$ 也成立． ▌

现在我们有下列定理．

定理 12.10

假设 x^* 和 μ^* 满足定理 2.8 给出的、x^* 是 (PE) 的严格局部极小值点的充分条件，则存在一个数 $c^*>0$，使得对所有 $c\geqslant c^*$，点 x^* 是 $M(x,\ \mu^*)$ 的一个局部无约束极小值点．反之，若 $h_j(x^0)=0$，$j=1,\ \cdots,\ p$，并且 x^0 是对某个 μ^0 的 $M(x,\ \mu^0)$ 的无约束极小值点，则 x^0 是(PE)的解．

证明由前面的结果可得出，留给读者去完成．

从上一定理能够看出，若 Lagrange 乘子 μ_j^* 的值和一个充分大的常数 c 可以得到，则 M 的一个无约束极小值点也成为(PE)的一个最优解．当然，困难在于决定 μ_j 和 c 的正确值．

Hestenes 的**乘子法**由每次迭代中修改 Lagrange 乘子 μ_j 有时也修改常数 c 所组成．假设我们选了一个充分大的 c 值，并设在第 k 次迭代中 Lagrange 乘子向量有某个估计 μ^k．我们求 $M(x,\ \mu^k)$的极小，并令最优解是 x^{k*}，注意到

$$\nabla_x M(x^{k*},\ \mu^*)=\nabla f(x^{k*})-\sum_{j=1}^{p}[\mu_j^k-ch_j(x^{k*})]\nabla h_j(x^{k*})=0.\tag{12.188}$$

Lagrange 乘子用以下公式修改：

$$\mu_j^{k+1}=\mu_j^k-ch_j(x^{k*}),\quad j=1,\ \cdots,\ p,\tag{12.189}$$

并进行下一次迭代，在这迭代中求 $M(x,\ \mu^{k+1})$的极小．

例 12.5.1

考虑二次规划

$$\min f(x)=\frac{1}{2}x^TQx+ab^Tx,\quad x\in R^n\tag{12.190}$$

受限制于

$$h(x)=\sum_{i=1}^{n} b_i x_i=0, \tag{12.191}$$

其中 Q 是非异对称阵，对所有满足(12.191)的 $x\neq 0$ 有 $x^TQx>0$. 向量 $b\neq 0$ 和数 $\alpha>0$ 是给定的. 读者能够验证，最优解是 $x^*=0$, 而 $\mu^*=\alpha$ 是相应的 Lagrange 乘子. 假设选择常数 c, 使对一切 $x\neq 0$ 有

$$x^TQx+c(b^Tx)^2>0, \tag{12.192}$$

则

$$M(x,\ \mu^k)=\frac{1}{2}x^TQx+\alpha b^Tx-\mu^k b^Tx+\frac{1}{2}c(b^Tx)^2, \tag{12.193}$$

M 的无约束极小值点 x^{k*} 必定满足方程

$$Qx^{k*}+\alpha b-(\mu^k-cb^Tx^{k*})b=0, \tag{12.194}$$

从上式我们得到

$$x^{k*}=\frac{(\mu^k-\alpha)Q^{-1}b}{1+cb^TQ^{-1}b}. \tag{12.195}$$

若从 $\mu^0=0$ 出发，则可证明

$$x^{k*}=\frac{-\alpha Q^{-1}b}{(1+cb^TQ^{-1}b)^{k+1}}, \tag{12.196}$$

从而，若

$$|1+cb^TQ^{-1}b|>1, \tag{12.197}$$

则序列 $\{x^{k*}\}$ 收敛于 $x^*=0$. ▐

Powell[48] 也独立地提出了等式约束问题的乘子法，十分类似于 Hestenes 的工作. 他考虑了下列双参数增广函数的无约束极小化：

$$M(x,\ \sigma,\ \alpha)=f(x)+\sum_{j=1}^{p}\sigma_j[h_j(x)+\alpha_j]^2. \tag{12.198}$$

注意，取 $\sigma_j=c/2$ 和 $\alpha_j=-\mu_j/c$, 则除了与 x 无关的一项之外，我们再次得到由 (12.179) 给出的 Hestenes 的增广函数. Powell 乘子法基于这样的结果：如果 $x^*(\sigma,\ \alpha)$ 是(12.198)对于参数 σ 和 α 的某组值的一个无约束极小值点，则 $x^*(\sigma,\ \alpha)$ 是 f 受限制于下列约束的极小值点：

$$h_j(x)=h_j[x^*(\sigma, \alpha)], \quad j=1, \cdots, p. \tag{12.199}$$

这就得出，参数 σ 和 α 必须用某种迭代方式调节，以使

$$\lim_{k\to\infty}\{h_j[x(\sigma^k, \alpha^k)]\}=0, \quad j=1, \cdots, p. \tag{12.200}$$

实际上，Powell 提议选取充分大的 σ_j, $j=1, \cdots, p$, 保持它们固定，并调节 α_j 使

$$\alpha_j^{k+1}=\alpha_j^k+h_j[x(\sigma, \alpha^k)], \quad j=1, \cdots, p. \tag{12.201}$$

继续校正参数 α_j, 直到 $\max|h_j(x(\sigma, \alpha))|$ 未能收敛或收敛于零的速率太慢，这时增加 σ_j 到较大的值.

这个方法对于(PE)的解的收敛性能够在比较温和的条件下证明. 我们介绍 Powell 的下列收敛性结果，其证明可在[48]中找到.

定理 12.11

假设 $x^*\in R^n$ 是(PE)的唯一解，并且对 x^* 的每个邻域 $N_\delta(x^*)$，存在 $\varepsilon(\delta)>0$，使得

$$X_\alpha=\{x:h_j(x)=\alpha_j, \quad j=1, \cdots, p\}\neq\emptyset, \tag{12.202}$$

其中每个 α_j 满足

$$|\alpha_j|\leqslant\varepsilon(\delta), \quad j=1, \cdots, p. \tag{12.203}$$

再假设 f, h_1, $\cdots$, h_p 是连续函数，并存在一自然数 K，使得对每个 $k\geqslant K$, $M(x, \sigma^k, \alpha^k)$的无约束极小值点落在一紧致集中. 则这些极小值点的序列收敛于 x^*.

这个方法收敛性的附带结果表明收敛速率是线性的. 然而，最近已表明这收敛速率能够改进. 类似于上述两种方法的另一种乘子法，也已由 Haarhoff、Buys[24] 提出. 又，Hestenes 乘子法已由 Miele 及其合作者[39,40,41,42]从计算角度予以广泛研究.

解等式约束非线性规划的乘子法由下列作者拓广到不等式约束问题：Bertsekas[3], Buys[5], Kort-Bertsekas[27,28], Pierre[45], Rockafellar[49,50]. 我们在这里简要叙述 Rockafellar 的工作.

假设要求解由(12.3)至(12.5)给出的等式-不等式约束问题(P). 定义

$$\theta(t)=\max\{t,\ 0\}.\tag{12.204}$$

那末对任何正数c,我们有

$$\frac{1}{4c}\{[\theta(\lambda_i-2cg_i(x))]^2-(\lambda_i)^2\}$$

$$=\begin{cases}c(g_i(x))^2-\lambda g_i(x), & g_i(x)\leqslant\dfrac{\lambda_i}{2c},\\ -\dfrac{(\lambda_i)^2}{4c}, & g_i(x)\geqslant\dfrac{\lambda_i}{2c},\end{cases}\tag{12.205}$$

并且增广 Lagrange 式 $M(x,\ \lambda,\ \mu)$可以定义为

$$M(x,\ \lambda,\ \mu)=f(x)+\frac{1}{4c}\sum_{i=1}^{m}\{[\theta(\lambda_i-2cg_i(x))]^2-(\lambda_i)^2\}$$
$$-\sum_{j=1}^{p}\mu_jh_j(x)+\frac{c}{2}\sum_{j=1}^{p}[h_j(x)]^2.\tag{12.206}$$

在练习 12.L 中,要求读者从先前的公式导出(12.206). 注意,与第 3 章的结果不同,出现在(12.206)中并相应于(P)的不等式约束的 Lagrange 乘子并不要求非负. Hestenes 和 Powell 的乘子法现在能够很容易地推广,对一个给定的 $c>0$ 和乘子$(\lambda^k,\ \mu^k)$,令 x^{k*} 是无约束最优化问题

$$\min_x M(x,\ \lambda^k,\ \mu^k)\tag{12.207}$$

的最优解. 再次注意到

$$\nabla_x M(x^{k*},\ \lambda^k,\ \mu^k)=\nabla f(x^{k*})-\sum_{i=1}^{m}\theta(\lambda_i^k-2cg_i(x^{k*}))\nabla g_i(x^{k*})$$
$$-\sum_{j=1}^{p}(\mu_j-ch_j(x^{k*}))\nabla h_j(x^{k*})=0.\tag{12.208}$$

所以取

$$\lambda_i^{k+1}=\theta(\lambda_i^k-2cg_i(x^{k*})),\quad i=1,\ \cdots,\ m,\tag{12.209}$$
$$\mu_j^{k+1}=\mu_j^k-ch_j(x^{k*}),\qquad j=1,\ \cdots,\ p\tag{12.210}$$

是合理的.

这个方法对于(P)的最优解的收敛性,能在类似于定理 12.11 的那些条件下证明. 此外,Rockafellar 还导出了收敛于凸规划问题(CP)一个解的另一种过程,(CP)由(4.158)至(4.160)给出,并

满足强相容性条件 (Slater 条件). 这个过程在每次迭代中不要求 M 的一个精确的极小值点，并且 c 的值能够保持不变. 关于细节，读者可参看[49].

至今所述的方法在计算上可能的缺点，是增广 Lagrange 式 M 通常只一次可微. 在本节余下部分，我们将统一、推广并在可能时改进这里谈及的工作. 这种统一做法的目的，是把一个一般的不等式-等式约束问题的最优解，与一类广义 Lagrange 式的相应无约束逗留点——甚至更好是这些函数的无约束极小值点——联系起来. 这个论题的下列结果属于 Mangasarian[37].

再次考虑由(12.3)至(12.5)给出的一般非线性规划(P). 定义广义 Lagrange 式

$$GL(x,\ y,\ w)=f(x)-\sum_{i=1}^{m}\phi(g_i(x),\ y_i)-\sum_{j=1}^{p}\psi(h_j(x),\ w_j),\tag{12.211}$$

其中 $\phi(z_1,\ z_2)$ 和 $\psi(z_1,\ z_2)$ 是定义在 R^2 上的实可微函数. 我们列出 ϕ 和 ψ 应满足的某些条件:

1. 方程

$$\frac{\partial\phi(0,\ \eta)}{\partial z_1}=\beta,\qquad \frac{\partial\phi(0,\ \eta)}{\partial z_2}=0\tag{12.212}$$

对每个 $\beta\geqslant0$ 有一个解 η; 方程

$$\frac{\partial\phi(\xi,\ \eta)}{\partial z_1}=0,\qquad \frac{\partial\phi(\xi,\ \eta)}{\partial z_2}=0\tag{12.213}$$

对每个 $\xi>0$ 有一个解 η.

2. 方程

$$\frac{\partial\psi(0,\ \eta)}{\partial z_1}=\beta,\qquad \frac{\partial\psi(0,\ \eta)}{\partial z_2}=0\tag{12.214}$$

对每个 β 有一个解 η.

3. 关系式

$$\frac{\partial\phi(0,\ \eta)}{\partial z_1}\geqslant0,\qquad \frac{\partial\phi(\xi,\ 0)}{\partial z_1}=0\tag{12.215}$$

对 $\xi>0$ 满足，并且，若

$$\frac{\partial\phi(\xi,\ \eta)}{\partial z_2}=0, \tag{12.216}$$

则 $\xi\geqslant0$ 和 $\xi\eta=0$.

4. 若

$$\frac{\partial\psi(\xi,\ \eta)}{\partial z_2}=0, \tag{12.217}$$

则 $\xi=0$.

建议读者在阅读下面结果之前，复习一下在定理 3.8 中提出的数学规划最优性的 Kuhn-Tucker 必要条件. 我们有下列定理.

定理 12.12

设 $f,\ g_1,\cdots,\ g_m,\ h_1,\ \cdots,\ h_p$ 在 x^* 可微. 若 x^* 是使条件(3.71)成立的 (P) 的一个局部解或整体解，或者若 x^* 和某个 $\lambda^*\in R^m$, $\mu^*\in R^p$ 满足 (3.72) 至 (3.74)，亦即关于 (P) 的最优性的广义 Kuhn-Tucker 必要条件成立，则 x^* 和某个 $y^*\in R^m$、$w^*\in R^p$ 成为 GL 的逗留点. 也就是说，倘若上面的条件(1)和(2)满足，则有

$$\nabla GL(x^*,\ y^*,\ w^*)=0. \tag{12.218}$$

反之，若 $(x^*,\ y^*,\ w^*)$ 满足(12.218)，其中

$$\lambda_i^*=\frac{\partial\phi(g_i(x^*),\ y_i^*)}{\partial z_1},\quad i=1,\ \cdots,\ m, \tag{12.219}$$

$$\mu_j^*=\frac{\partial\psi(h_j(x^*),w_j^*)}{\partial z_1},\quad j=1,\ \cdots,\ p, \tag{12.220}$$

则当条件(3)和(4)成立时，向量 $(x^*,\ \lambda^*,\ \mu^*)$ 满足(P)的最优性的广义 Kuhn-Tucker 必要条件.

【证明】 若 x^* 是(P)的一个局部或整体解且(3.71)成立，则 Kuhn-Tucker 条件(3.72)至(3.74)对 x^* 和某个 λ^*、μ^* 满足. 令 ϕ 和 ψ 满足条件(1)和(2)，$I(x^*)$ 是积极约束集，并定义 y^*、w^* 为

$$\frac{\partial\phi(0,\ y_i^*)}{\partial z_1}=\lambda_i^*,\qquad \frac{\partial\phi(0,\ y_i^*)}{\partial z_2}=0,\qquad i\in I(x^*), \tag{12.221}$$

$$\frac{\partial\psi(g_i(x^*),\ y_i^*)}{\partial z_1}=0,\quad \frac{\partial\phi(g_i(x^*),y_i^*)}{\partial z_2}=0,\quad i\bar{\in} I(x^*), \tag{12.222}$$

$$\frac{\partial\psi(0,\ w_j^*)}{\partial z_1}=\mu_j^*,\quad \frac{\partial\psi(0,\ w_j^*)}{\partial z_2}=0,\quad j=1,\ \cdots,\ p. \tag{12.223}$$

于是，由(3.72)至(3.74)以及(12.221)至(12.223)可导出

$$\nabla_x GL(x^*,\ y^*,\ w^*)=\nabla f(x^*)-\sum_{i=1}^{m}\frac{\partial\phi(g_i(x^*),\ y_i^*)}{\partial z_1}\nabla g_i(x^*) -\sum_{j=1}^{p}\frac{\partial\psi(h_j(x^*),\ w_j^*)}{\partial z_1}\nabla h_j(x^*)=0, \tag{12.224}$$

$$\frac{\partial GL(x^*,\ y^*,\ w^*)}{\partial y_i}=-\frac{\partial\phi(g_i(x^*),\ y_i^*)}{\partial z_2}=0,\quad i=1,\ \cdots,\ m, \tag{12.225}$$

$$\frac{\partial GL(x^*,\ y^*,\ w^*)}{\partial w_j}=-\frac{\partial\psi(h_j(x^*),w_j^*)}{\partial z_2}=0,\quad j=1,\ \cdots,\ p. \tag{12.226}$$

因此

$$\nabla GL(x^*,\ y^*,\ w^*)=0. \tag{12.227}$$

反之，假设$(x^*,\ y^*,\ w^*)$满足(12.218)，并设条件(3)和(4)成立. 用(12.219)和(12.220)分别定义λ^*和μ^*，从(12.224)我们得到

$$\nabla f(x^*)-\sum_{i=1}^{m}\lambda_i^*\nabla g_i(x^*)-\sum_{j=1}^{p}\mu_j^*\nabla h_j(x^*)=0. \tag{12.228}$$

从条件(3)中的(12.215)和(12.219)就有

$$\lambda_i^*\geqslant 0,\quad i\in I(x^*). \tag{12.229}$$

从条件(3)中的(12.216)和(12.225)，我们得到

$$g_i(x^*)\geqslant 0,\quad i=1,\ \cdots,\ m, \tag{12.230}$$

$$y_i^* g_i(x^*)=0,\quad i=1,\ \cdots,\ m. \tag{12.231}$$

从(12.215)、(12.219)和(12.231)，我们有

$$\lambda_i^*=\frac{\partial\phi(g_i(x^*),\ y_i^*)}{\partial z_1}=\frac{\partial\psi(g_i(x^*),\ 0)}{\partial z_1}=0,\quad i\overline{\in} I(x^*). \tag{12.232}$$

最后，从(12.217)和(12.226)，我们有

$$h_j(x^*)=0,\quad j=1,\ \cdots,\ p. \tag{12.233}$$

把(12.229), (12.230)和(12.232)综合为下列等价的表示式:

$$\lambda_i^* g_i(x^*)=0, \quad i=1, \cdots, m, \tag{12.234}$$

$$\lambda^*\geqslant 0, \tag{12.235}$$

我们断定(12.228), (12.230)以及(12.233)至(12.235)是定理3.8给出的Kuhn-Tucker条件. ▌

若进一步的条件加于ϕ和ψ, 则可保证对参数α的充分大的值, 由定理第一部分给出的、GL的一个逗留点(x^*, y^*, w^*), 对应于$GL(x, y^*, w^*)$的一个严格无约束局部极小值点. 从而在恰当惩罚函数和增广Lagrange式中的前述结果被推广了.

考察函数ϕ和ψ的某些例子, 它们满足上一定理所要求的条件(1)至(4), 以及保证GL在(x^*, y^*, w^*)的Hesse阵为正定的条件. 定义

$$\Gamma(t)_+=\begin{cases}\Gamma(t), & t\geqslant 0,\\ 0, & t<0,\end{cases} \tag{12.236}$$

并令

$$\phi(\xi, \eta)=\Gamma(\eta)-\Gamma(\eta-\alpha\xi)_+, \tag{12.237}$$

$$\psi(\xi, \eta)=\Gamma(\eta)-\Gamma(\eta-\alpha\xi), \tag{12.238}$$

其中Γ是R上一个可微实函数, $\alpha\in R$是一个参数. 函数Γ本身必须满足一定的条件: (a)导数Γ'是把R映射到R上的严格递增函数, $\Gamma'(t)_+$把$[0, \infty)$映射到自身, 且$\Gamma'(0)=0$; (b) Γ是R上使$\Gamma(0)=0$的一个非负凸函数; (c) Γ是R上二次可微函数, 对$t\neq 0$有$\Gamma''(t)>0$, 并且$\Gamma''(0)=0$. 满足条件(a)、(b)和(c)的函数Γ的例子是

$$\Gamma_1(t)=\frac{1}{\alpha\beta}(t)^{\beta}, \quad \alpha>0, \quad \beta=4, 6, \cdots, \tag{12.239}$$

$$\Gamma_2(t)=\frac{1}{2}[e^t+e^{-t}-(t)^2]-1, \tag{12.240}$$

$$\Gamma_3(t)=\frac{1}{2}\left[\frac{1}{2}(e^t+e^{-t})-1\right]^2. \tag{12.241}$$

对应于这些函数, 我们得到

$$\phi_1(\xi,\ \eta)=\frac{1}{\alpha\beta}[(\eta)^\beta-(\eta-\alpha\xi)_+^\beta],\quad \alpha>0,\quad \beta=4,\ 6,\ \cdots,\tag{12.242}$$

$$\psi_1(\xi,\ \eta)=\frac{1}{\alpha\beta}[(\eta)^\beta-(\eta-\alpha\xi)^\beta],\quad \alpha>0,\quad \beta=4,\ 6,\ \cdots\tag{12.243}$$

等等. 令 $\beta=4$,我们能把规划(P)的广义 Lagrange 式写为

$$GL_1(x,\ y,\ w)=f(x)-\frac{1}{4\alpha}\sum_{i=1}^{m}[(y_i)^4-(y_i-\alpha g_i(x))_+^4]$$
$$-\frac{1}{4\alpha}\sum_{j=1}^{p}[(w_j)^4-(w_j-\alpha h_j(x))^4].\tag{12.244}$$

满足(a)、(b)、(c)(除条件 $\Gamma''(0)=0$ 以外)的函数的一个例子给出为

$$\Gamma_4(t)=\frac{1}{2\alpha}(t)^2,\quad \alpha>0,\tag{12.245}$$

从它我们得到

$$\phi_4(\xi,\ \eta)=\frac{1}{2\alpha}[(\eta)^2-(\eta-\alpha\xi)_+^2],\tag{12.246}$$

$$\psi_4(\xi,\ \eta)=\frac{1}{2\alpha}[(\eta)^2-(\eta-\alpha\xi)^2],\tag{12.247}$$

和

$$GL_4(x,\ y,\ w)=f(x)-\frac{1}{2\alpha}\sum_{i=1}^{m}[(y_i)^2-(y_i-\alpha g_i(x))_+^2]$$
$$-\frac{1}{2\alpha}\sum_{j=1}^{p}[(w_j)^2-(w_j-\alpha h_j(x))^2].\tag{12.248}$$

令 $c=2$、$\lambda_i=y_i$ 和 $\mu_j=w_j$, 读者容易验证 GL_4 恰好是(12.206)给出的 Rockafellar 的增广 Lagrange 式,它不是处处二次可微的,这与基于(a)、(b)和(c)的所有条件的函数 Γ 的广义 Lagrange 式是不相同的.

Mangasarian[37] 也得到了基于广义 Lagrange 式的对偶结果以及求 GL 极小的超线性(或二次)收敛算法. 这个算法本质上是用 Newton 法去解一族非线性方程 $\nabla GL(x,\ y,\ w)=0$. 在有关的发展中, Mangasarian 把乘子法扩充到不等式约束和更一般的

Lagrange 式中去，并得到线性收敛的算法.

12.6 惩罚函数法的若干计算背景

在头三节中所引入的惩罚函数法，由一系列无约束极小化所组成，广泛地被用于求解非线性最优化问题. 第 12.4 节和 12.5 节的方法更为新近，所以计算经验十分有限. 然而我们相信，它们能够十分有效，并且可能成为主要的全能工具之一，用以获得高度非线性问题的数值解.

惩罚函数法，即以归并约束函数到无约束问题的目标函数中去的方式把约束问题化为无约束问题的方法，预期对非线性约束函数能特别有效. 因为在后几章中我们将看到，线性约束能用各种其他方法有效地处理，所以把它们归并到惩罚函数中不是特别有用. 外部惩罚法和内部惩罚法存在着简单的修改[15]，用一种混合的"非约束-简单的"(unconstrained-simplicial) 算法单独地处理线性约束. 解非线性规划的真正挑战在于处理非线性约束. 对这样的问题，惩罚函数具有某些肯定的优点，最重要的或许是，与其他直截了当地处理非线性约束的方法大不相同，惩罚函数法避免了沿能行集边界线移动或力图不越出边界这一花费时间且又十分艰巨的任务. 另一个优点是，在多数场合下，为了应用这个方法，无须对出现于问题中的函数作一些特殊的假定如凸性或广义凸性之类. 后面我们将看到，有些其他算法只对特殊的函数族定义，这样就不能普遍使用[38]. 虽然许多算法收敛于问题的解，它满足定理 3.8 给出的 Kuhn-Tucker 一阶必要条件，然而序列无约束极小化方法一般收敛于满足定理 3.10 的二阶必要条件的点. 因此，在这些点上也期望能满足二阶充分条件. 能够证明，在二阶条件成立的点的近旁，序列惩罚方法的迭代常常由严格凸函数的无约束极小化组成，见 McCormick[38] 和 Polyak[47].

用无约束极小化有效地求解约束问题，强烈地依赖于实行无约束最优化的数值方法. 最有效的无约束方法是第 10 章和第 11 章中讨论的一批，这些方法利用由求极小的函数的梯度以及有时

是 Hesse 阵的值的计算所提供的信息．有讽刺意味的是，这种方法对于序列惩罚方法的主要缺点而言是最脆弱的：随着惩罚方法的参数趋于它的极限，惩罚函数的 Hesse 阵逐渐变得更病态．例如，考察一个如同(12.82)那样的混合惩罚函数 $H(x, \rho)$，它是求解具有等式和不等式约束问题(P)给出的．令

$$H(x, \rho^k)=f(x)-\rho^k\sum_{i=1}^{m}\log g_i(x)+\frac{1}{\rho^k}\sum_{j=1}^{p}(h_j(x))^2, \tag{12.249}$$

其中 $\rho^k>0$．并且当$\{\rho^k\}\to 0$时实行对 H 的一系列无约束极小化．H 的 Hesse 阵由下式给出：

$$\begin{aligned}\nabla^2 H(x, \rho^k)=\nabla^2 f(x)-\sum_{i=1}^{m}\frac{\rho^k}{g_i(x)}\nabla^2 g_i(x)+\sum_{j=1}^{p}\frac{2h_j(x)}{\rho^k}\nabla^2 h_j(x)\\+\sum_{i=1}^{m}\frac{\rho^k}{(g_i(x))^2}\nabla g_i(x)(\nabla g_i(x))^T\\+\sum_{j=1}^{p}\frac{2}{\rho^k}\nabla h_j(x)(\nabla h_j(x))^T.\end{aligned} \tag{12.250}$$

能够证明[15]，序列 $\{\rho^k/g_i(x^{k*})\}$ 和 $\{2h_j(x^{k*})/\rho^k\}$ 在温和假定下分别收敛于相应的 Lagrange 乘子 λ_i^* 和 μ_j^*．这样，$\nabla^2 H$ 的前三项收敛于 Lagrange 式的 Hesse 阵，它有有限的特征值．然而，余下两项包含序列$\{\rho^k/g_i(x))^2\}$ 和 $\{2/\rho^k\}$，它们在多数情况下无界地递增从而造成病态．这个问题近来已受到一些注意，提出了惩罚方法的几种修改．读者可参考 Fletcher、McCann[20]，Lootsma[34, 36]，以及 Murray[43]．由 Lootsma[34] 作的问题的分析表明，对于由(12.249)给出、利用(12.51)中函数 q 的函数 $H(x, \rho^k)$，比起利用(12.52)或(12.53)的惩罚函数来说，Hesse 阵的病态更加快．另一方面，带有(12.51)中函数 q 的收敛性通常比用其他两个函数快．一个有关的困难是，惩罚函数的 Hesse 逆阵趋于奇异，这主要出现在用变尺度算法求解序列无约束极小化的时候．我们期望，借助于转向恰当惩罚法或乘子法，惩罚函数法的那些与 Hesse 阵联系的大多数缺点将会消除．

Lootsma[35] 检验了前几章叙述的几种无约束最优化方法，在

应用于求惩罚函数的极小时的执行情况．在有限几个试验问题的基础上，他推荐利用 BFS 方法或修改的 Newton 方法[15]．惩罚方法的一种特殊的无约束技术以及障碍函数的一种线性搜索已分别由 Lasdon[30] 和 Lasdon、Fox、Ratner[32] 提出．

在惩罚函数法中一个有趣而有用的特点，是借助于外推来加速该方法的收敛性，如 Fiacco、McCormik[16] 所建议．其想法是在建立最优点序列 $\{x^{k*}\}$ 时，利用这个序列中已有点的级数展开去估计相继点．所得到的这个信息能用来开始带有下一个惩罚参数值的无约束最优化．可以想象，这种加速技术对其他类型的算法也是有用的，但这个论题还没有受到适当的注意．

练　习

12.A． 画一个求解问题 (P) 的外部惩罚函数法的框图．假定无约束最优化是执行：(a) 9.5 节的 Powell 方法，或 (b) BFS 方法那样的变尺度算法．讨论两种情况下计算上可能的优缺点，并列举实现这两种方法时预期会产生的困难．又，倘若不等式约束问题的最优解落在能行集内部，讨论外部惩罚函数法的性态，并用数值例子说明这个性态．

12.B． 试证由

$$X_1=\{x:x\in R^2,\ (x_2)^2-(x_1x_2)^2+(x_2(x_2-1))^2\geqslant 0\} \qquad (12.251)$$

给出的集合 X_1，并不满足 12.2 节中对(PI)所假定的正则性条件．

令

$$X_2=\{x:x\in R^n,\ g(x)\geqslant 0\}, \qquad (12.252)$$

求一个实函数 g，使 X_2 的内部 X_2^0 非空，但对每个 $x\in X_2^0$ 有 $g(x)=0$．

12.C． 证明无参数外部方法在 12.1 节所述条件下的收敛性．

12.D． 用本章讨论的惩罚函数法解下列约束问题[15,40]．利用前几章中发展的无约束极小化的数值方法，这些方法在你的计算机中是可用的．

(a)
$$\min\{(x_1)^3-6(x_1)^2+11x_1+x_3\} \qquad (12.253)$$

受限制于

$$(x_1)^2+(x_2)^2-(x_3)^2\leqslant 0, \qquad (12.254)$$

$$(x_1)^2+(x_2)^2+(x_3)^2\geqslant 4, \qquad (12.255)$$

$$x_3 \leqslant 5, \tag{12.256}$$

$$x_1 \geqslant 0, \quad x_2 \geqslant 0, \quad x_3 \geqslant 0. \tag{12.257}$$

一个最优解在 $x^* = (0, \sqrt{2}, \sqrt{2})$.

(b)
$$\min\{(x_1 - x_2)^2 + (x_2 - x_3)^4\} \tag{12.258}$$

受限制于

$$x_1 + x_1(x_2)^2 + (x_3)^4 = 3. \tag{12.259}$$

一个最优解在 $x^* = (1, 1, 1)$.

12.E. 按照利用(12.10)和(12.14)的外部惩罚函数法求解例12.5.1中的二次规划. 令 $1/\rho^k = (k+1)/2$ 和 $\beta = 2$, 证明极小值点序列 $\{x^{k*}\}$ 由下式确定:

$$x^{k*} = \frac{-\alpha Q^{-1}b}{1 + (k+1)b^T Q^{-1} b}. \tag{12.260}$$

把这个序列的收敛性与例题中给出的乘子法的收敛性作比较.

12.F. 证明关于恰当惩罚函数的下列论断: 设对 $\beta > 0$ 和一切 $\lambda \in A \subset R^m$, 成立着

$$\gamma_i(\beta, \lambda) = 0, \quad i = 1, \cdots, m. \tag{12.261}$$

假设 $\lambda \in A$, 又 $\hat{x}$ 关于(PI)是能行的, 且 $\hat{x}$ 是(12.132)的一个局部极小值点, 那末 $\hat{x}$ 也是问题(PI)的一个局部极小值点.

12.G. 试证, 在第2章和第3章叙述的最优性的Kuhn-Tucker必要条件, 对于等式约束规划(PE)来说, 在 $\hat{x}$ 是

$$P\nabla f(\hat{x}) = 0, \tag{12.262}$$

$$h(\hat{x}) = 0, \tag{12.263}$$

其中 P 由(12.154)给出. 给出(12.262)的一个几何解释, 并证明

$$\nabla\phi(\hat{x}) = 0, \tag{12.264}$$

其中 ϕ 由(12.155)定义.

12.H. 证明定理12.7.

12.I. 假设在例12.5.1中所述的二次规划中的 Q 是正定的, 用带有 $\sigma = 1$ 和 $\alpha^0 = 0$ 的Powell乘子法去解这个规划. 将得到收敛于 $x^* = 0$ 吗?

12.J. 证明定理12.10.

12.K. 一个等式-不等式约束的非线性规划, 诸如本章开始时提出的(P). 用引进松弛变量的方法, 能把它转变为等价的规划. 这样, 规划(P)能重写为

(P′)
$$\min f(x) \tag{12.265}$$

受限制于

$$g_i(x)-z_i=0,\quad i=1,\cdots,m,\tag{12.266}$$

$$h_j(x)=0,\quad j=1,\cdots,p,\tag{12.267}$$

$$z_i\geqslant 0,\quad i=1,\cdots,m.\tag{12.268}$$

用 Hestenes 乘子法求解这个规划，就是找一个 $x\in R^n$ 和 $z\in R^m$，$z\geqslant 0$，使下述函数极小化：

$$M(x,z,\lambda,\mu)=f(x)-\sum_{i=1}^{m}\lambda_i(g_i(x)-z_i)+\frac{c}{2}\sum_{i=1}^{m}(g_i(x)-z_i)^2$$

$$-\sum_{j=1}^{p}\mu_j h_j(x)+\frac{c}{2}\sum_{j=1}^{p}(h_j(x))^2.\tag{12.269}$$

注意，相对于 $z\geqslant 0$ 的极小化能够明显地作出，从而得到 $M(x,\lambda,\mu)$ 的公式(12.206)。

12.L. 假设(SP)是 5.2 节定义的一个标准原有凸规划，即

$$\text{(SP)}\qquad \min f(x)\tag{12.270}$$

受限制于

$$g_i(x)\geqslant 0,\quad i=1,\cdots,m.\tag{12.271}$$

(a) 从(5.120)和下列扰动问题导出对应于(SP)的形如(12.206)的增广 Lagrange 式 $M(x,\lambda)$：

$$\text{(QSP)}\qquad \min f(x)+c\sum_{i=1}^{m}(w_i)^2,\tag{12.272}$$

受限制于

$$g_i(x)\geqslant w_i,\quad i=1,\cdots,m,\tag{12.273}$$

其中 $c>0$ 是参数。

(b) 证明对应于(QSP)的扰动函数 Φ_Q 为

$$\Phi_Q(w)=\Phi_P(w)+c\sum_{i=1}^{m}(w_i)^2,\tag{12.274}$$

其中 Φ_P 是对应于由(5.111)和(5.112)定义的(PSP)的扰动函数。

12.M. 用例 3.1.4 中提出问题说明定理 12.12. 取 $x^*=(0,0)$，并选择 $\alpha=1$，$y^*=\frac{1}{2}$，$w^*=0$. 按(12.244)定义 GL. $\left(0,0,\frac{1}{2},0\right)$是广义 Lagrange 式当 $\alpha\geqslant 1$ 时的一个极小值点吗?

12.N. 证明：由(12.239)至(12.241)给出的 Γ_1、Γ_2 和 Γ_3，必定如断言的那样满足条件(a)，(b)和(c)。

参考文献

1. ARROW, K. J., F. J. GOULD, and S. M. HOWE, "A General Saddle Point Result for Constrained Optimization," *Math. Prog.*, **5**, 225–234 (1973).
2. BARTLE, R. G., *The Elements of Real Analysis*, 2nd ed., John Wiley & Sons, New York, 1976.
3. BERTSEKAS, D. P., "On Penalty and Multiplier Methods for Constrained Minimization," Engineering-Economic Systems Department Working Paper, Stanford University, August 1973, *SIAM J. Control*, to appear.
4. BRACKEN J., and G. P. McCORMICK, *Selected Applications of Nonlinear Programming*, John Wiley & Sons, New York, 1968.
5. BUYS, J. D., "Dual Algorithms for Constrained Optimization Problems," Doctoral dissertation, University of Leiden, The Netherlands, 1972.
6. CARROLL, C. W., "The Created Response Surface Technique for Optimizing Nonlinear Restrained Systems," *Operations Research*, **9**, 169–184 (1961).
7. CONN, A. R., "Constrained Optimization Using a Nondifferentiable Penalty Function," *SIAM J. Numer. Anal.*, **10**, 760–784 (1973).
8. CONN, A. R., and T. PIETRZYKOWSKI, "A Penalty Function Method Converging Directly to a Constrained Optimum," Research Report 73-11, Department of Combinatorics and Optimization, University of Waterloo, Ontario, Canada, May 1973, *SIAM J. Numer. Anal.*, to appear.
9. COURANT, R., "Variational Methods for the Solution of Problems of Equilibrium and Vibrations," *Bull. Am. Math. Soc.*, **49**, 1–23 (1943).
10. EVANS, J. P., and F. J. GOULD, "An Existence Theorem for Penalty Function Theory," *SIAM J. Control*, **12**, 509–516 (1974).
11. EVANS, J. P., F. J. GOULD, and J. W. TOLLE, "Exact Penalty Functions in Nonlinear Programming," *Math. Prog.*, **4**, 72–97 (1973).
12. FIACCO, A. V., and G. P. McCORMICK, "The Sequential Unconstrained Minimization Technique for Nonlinear Programming: A Primal-Dual Method," *Management Science*, **10**, 360–366 (1964).
13. FIACCO, A. V., and G. P. McCORMICK, "Extensions of SUMT for Nonlinear Programming: Equality Constraints and Extrapolation," *Management Science*, **12**, 816–829 (1966).
14. FIACCO, A. V., and G. P. McCORMICK, "The Sequential Unconstrained Minimization Technique without Parameters," *Operations Research*, **15**, 820–827 (1967).
15. FIACCO, A. V., and G. P. McCORMICK, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, John Wiley & Sons, New York, 1968.
16. FLETCHER, R., "A Class of Methods for Nonlinear Programming with Termination and Convergence Properties," in *Integer and Nonlinear Programming*, J. Abadie (Ed.), North-Holland Publishing Co., Amsterdam, 1970.

17. FLETCHER, R., "A Class of Methods for Nonlinear Programming III, Rates of Convergence," in *Numerical Methods for Non-linear Optimization*, F. A. Lootsma (Ed.), Academic Press, London, 1972.

18. FLETCHER, R., "An Exact Penalty Function for Nonlinear Programming with Inequalities," *Math. Prog.*, **5**, 129–150 (1973).

19. FLETCHER, R., and S. A. LILL, "A Class of Methods for Nonlinear Programming II, Computational Experience," in *Nonlinear Programming*, J. B. Rosen, O. L. Mangasarian, and K. Ritter (Eds.), Academic Press, New York, 1970.

20. FLETCHER, R., and A. P. McCANN, "Acceleration Techniques for Nonlinear Programming," in *Optimization*, R. Fletcher (Ed.), Academic Press, London, 1969.

21. FOX, R. L., "Structural and Mechanical Design Optimization," in *Optimization and Design*, M. Avriel, M. J. Rijckaert, and D. J. Wilde (Eds.), Prentice-Hall, Englewood Cliffs, N.J., 1973.

22. FRISCH, K. R., "The Logarithmic Potential Method of Convex Programming," University Institute of Economics, Oslo, Norway, Memorandum, May 13, 1955.

23. GOULD, F. J., "A Class of Inside-out Algorithms for General Programs," *Management Science*, **16**, 350–356 (1970).

24. HAARHOFF, P. C., and J. D. BUYS, "A New Method for the Optimization of a Nonlinear Function Subject to Nonlinear Constraints," *Computer J.*, **13**, 178–184 (1970).

25. HESTENES, M. R., "Multiplier and Gradient Methods," *J. Optimization Theory & Appl.*, **4**, 303–320 (1969).

26. HUARD, P., "Resolution of Mathematical Programming with Nonlinear Constraints by the Method of Centres," in *Nonlinear Programming*, J. Abadie (Ed.), North-Holland Publishing Co., Amsterdam, 1967.

27. KORT, B. W., and D. P. BERTSEKAS, "A New Penalty Function Method for Constrained Minimization," in *Proceedings of 1972 IEEE Conference on Decision and Control*, New Orleans, December 1972.

28. KORT, B. W., and D. P. BERTSEKAS, "Combined Primal-Dual and Penalty Methods for Convex Programming," Engineering-Economic Systems Dept. Working Paper, Stanford University, August 1973, *SIAM J. Control*, to appear.

29. KOWALIK, J., M. R. OSBORNE, and D. M. RYAN, "A New Method for Constrained Optimization Problems," *Operations Research*, **17**, 973–983 (1969).

30. LASDON, L. S., "An Efficient Algorithm for Minimizing Barrier and Penalty Functions," *Math. Prog.*, **2**, 65–106 (1972).

31. LASDON, L. S., "Penalty and Barrier Function Design of Electrical Networks and Sonar Transducer Arrays," in *Optimization and Design*, M. Avriel, M. J. Rijckaert, and D. J. Wilde (Eds.), Prentice-Hall, Englewood Cliffs, N.J., 1973.

32. LASDON, L. S., R. L. FOX, and M. W. RATNER, "An Efficient One-Dimensional Search Procedure for Barrier Functions," *Math. Prog.*, **4**, 279–296 (1973).

33. Lootsma, F. A., "Constrained Optimization via Parameter-Free Penalty Functions," *Philips Res. Repts.*, **23**, 424–437 (1968).

34. Lootsma, F. A., "Hessian Matrices of Penalty Functions for Solving Constrained Optimization Problems," *Philips Res. Repts.*, **24**, 322–330 (1969).

35. Lootsma, F. A., "Penalty Function Performance of Several Unconstrained Minimization Techniques," *Philips Res. Repts.*, **27**, 358–385 (1972).

36. Lootsma, F. A., "A Survey of Methods for Solving Constrained Minimization Problems via Unconstrained Minimization," in *Numerical Methods for Nonlinear Optimization*, F. A. Lootsma (Ed.), Academic Press, London, 1972.

37. Mangasarian, O. L., "Unconstrained Lagrangians in Nonlinear Programming," *SIAM J. Control*, **13**, 772–791 (1975).

38. McCormick, G. P., "Penalty Function versus Non-penalty Function Methods for Constrained Nonlinear Programming Problems," *Math. Prog.*, **1**, 217–238 (1971).

39. Miele, A., E. E. Cragg, R. R. Iver, and A. V. Levy, "Use of the Augmented Penalty Function in Mathematical Programming Problems, Part 1," *J. Optimization & Theory & Appl.*, **8**, 115–130 (1971).

40. Miele, A., E. E. Cragg, and A. V. Levy, "Use of the Augmented Penalty Function in Mathematical Programming Problems, Part 2," *J. Optimization Theory & Appl.*, **8**, 131–153 (1971).

41. Miele, A., P. E. Moseley, and E. E. Cragg, "A Modification of the Method of Multipliers for Mathematical Programming Problems," in *Techniques of Optimization*, A. V. Balakrishnan (Ed.), Academic Press, New York, 1972.

42. Miele, A., P. E. Moseley, A. V. Levy, and G. M. Coggins, "On the Method of Multipliers for Mathematical Programming Problems," *J. Optimization Theory & Appl.*, **10**, 1–33 (1972).

43. Murray, W., "Analytical Expressions for the Eigenvalues and Eigenvectors of the Hessian Matrices of Barrier and Penalty Functions," *J. Optimization Theory & Appl.*, **7**, 189–196 (1971).

44. Noble, B., *Applied Linear Algebra*, Prentice-Hall, Englewood Cliffs, N.J., 1969.

45. Pierre, D. A., "Multiplier Algorithms for Nonlinear Programs," paper presented at the Eighth Int. Symp. on Math. Prog., Stanford University, Stanford, Ca., August 1973.

46. Pietrzykowski, T., "An Exact Potential Method for Constrained Maxima," *SIAM J. Numer. Anal.*, **6**, 299–304 (1969).

47. Polyak, B. T., "The Convergence Rate of the Penalty Function Method," *USSR Comp. Math. & Math. Phys.*, **11**, 1, 1–12 (1971).

48. Powell, M. J. D., "A Method for Nonlinear Constraints in Minimization Problems," in *Optimization*, R. Fletcher (Ed.), Academic Press, London, 1969.

49. Rockafellar, R. T., "The Multiplier Method of Hestenes and Powell Applied to Convex Programming," *J. Optimization Theory & Appl.*, **12**, 555–562 (1973).

50. ROCKAFELLAR, R. T., "A Dual Approach to Solving Nonlinear Programming Problems by Unconstrained Optimization," *Math. Prog.*, **5**, 354–373 (1973).

51. RUSH, B. C., J. BRACKEN, and G. P. McCORMICK, "A Nonlinear Programming Model for Launch Vehicle Design and Costing," *Operations Research*, **15**, 185–210 (1967).

52. ZANGWILL, W. I., "Non-linear Programming via Penalty Functions," *Management Science*, **13**, 344–358 (1967).

第 13 章

用无约束最优化方法的扩充求解约束问题

前一章我们把约束结合进目标函数，作间接的处理．与此相反，这里讨论直接处理约束的方法．现在给出的方法是为无约束最优化问题导出的方法的扩充，其中引进了考虑约束的额外步骤．这些步骤首先寻找一个导向约束极小值点的下降方向，然后执行一个限制在能行集中的线性搜索．下面将会表明，对线性约束问题，这些额外的步骤是不难完成的．如变尺度法等某些有效的无约束方法在出现线性约束时，也是非常有效的．因为这样的算法是成功的，普遍认为，即使应用惩罚函数法，线性约束仍应当直接进行处理，而不是被转化成不同形式的增广目标函数．关于一般非线性约束，没有这样的一致看法，但是可以看到，在目前的数值方法状况下，对这种类型的问题宁愿用惩罚方法．在直接方法中，跟踪积极非线性约束(如等式约束)的过程通常是一个困难而又费时的工作，应当避免．一种可能的选择是将非线性约束相继线性化，其结果一般就成为由求解一系列线性规划组成的一种方法．这种处理在第 14 章讨论．

第 13.1 节给出不用导数的、无约束最优化的经验方法对约束问题的简单推广．这些方法在无约束问题中的做法是在第 9 章中讨论的．梯度投影或“偏转”梯度法是 13.2 节的主题，它们是最速下降法和变尺度算法的扩充．在 13.3 节中有二次规划的一个算法，它说明了上述方法的一种特殊情形．对于一般函数在约束下的极小化问题的求解来说，能行方向法是首批设计的一个数值方法，对于线性约束情形的这一方法在 13.4 节中讲述．最后，以投影方向法和能行方向法直接处理非线性约束来结束本章，同时给出几个把非线性约束问题简化为线性约束问题的新方法．

13.1 经验方法的扩充

在9.1节至9.3节中介绍的各个经验方法都已被扩充，用来处理约束问题.我们以单纯形法的有约束形式开始讨论这种扩充.

设所考虑的问题是：求实函数 f 的极小值点 $x^* \in R^n$，受限制于不等式约束(已表示为记号方便的形式)：

$$l_j \leqslant x_j \leqslant u_j, \quad j=1, \cdots, m, \tag{13.1}$$

其中 $x=(x_1, \cdots, x_n)^T$ 是独立变量向量，且

$$x_{n+i}=g_i(x), \quad i=1, \cdots, m-n. \tag{13.2}$$

上界 u_j 和下界 l_j 假定是已知数或 x 的函数. Box[5] 把单纯形法扩充去解这问题，相应地，他把有约束单纯形法称为**复合形法**. 先是应用于 Spendley、Hext、Himsworth [38] 的方法，但也能应用于 Nelder、Mead[23] 的改进单纯形法.

假定已给一个满足(13.1)中所有 m 个约束的初始能行点 x^0. Box 提议，不是产生初始单纯形的其余 n 个点，而是产生 $N \geqslant n$ 个附加点，这里 N 的典型值为 $2n$. 这些点一次选一个，使得它们的第 j 个分量满足

$$x_j^k=l_j+r_j^k(u_j-l_j), \quad j=1, \cdots, n, \quad k=1, \cdots, N, \tag{13.3}$$

其中 r_j^k 是开区间 $(0, 1)$ 上均匀分布的伪随机数. 由(13.3)定义的点将满足(13.1)中前 n 个约束，但可能违反"隐式"约束即第 $n+1, \cdots, m$ 个约束. 在这种情形下，这些非能行的试验点就向已选定的能行点的形心缩回一半，直到最后所有 N 点都变成能行的.

其次决定函数值，并以 N 代替 n 算出式(9.1)至(9.4)的有关值. 如果从反映步骤得到的点 x^r 是能行的，则计算就顺着无约束单纯形法的步骤进行，直至遇到非能行点. 如果一点违反了(13.1)中前 n 个约束之一，则违反了的分量重新取值于适当的范围以给出进一步的试验点. 但如果某个"隐式"约束被违反，则向着进行反映的形心缩回一半以产生进一步的试验点，直至该点变成能行点.

Mitchell、Kaplan[21] 提出另一个不用随机数而获得初始复合形的过程．他们的初始复合形由 $2n+1$ 个点组成，通常比 Box 方法构造得的要大，这样，它就增加了机会以求得整体极小值点而不是局部极小值点．

在一个复合形中使用多于 $n+1$ 个点的原因在于，向约束的能行一侧收缩的结果，由 $n+1$ 个点组成的单纯形趋于崩溃，不再张成全空间了．结果真正的极小值点容易被遗漏．复合形法的一般执行情况类似于无约束时的做法．它是一个较慢的但可靠的方法，在解一个给定问题时，它的成功紧密地依赖于初始点，它的效率随着变数个数的增加而很快地衰减．

Glass、Cooper[13] 把 9.2 节的模式搜索法推广到约束最优化．在他们的方法中，搜索从一个能行点开始，一直继续至一个或多个约束已达到，因而在违反约束时已不可能再用勘探移动来继续这方法．这时，由另一个子程序借助于解某个线性规划问题来决定一个新方向，这线性规划问题是将目标函数与约束函数作线性近似而形成的．因为这个方法比复合形法或者下面将描述的方法复杂得多，同时，比本章后面讨论的几个更巧妙的算法效率低得多，所以它的应用不广泛．这个方法的细节和某些数值经验可以在[13]中找到．

如 9.3 节所指出的，Rosenbrock 的旋转方向法[37] 是为有约束和无约束极小化问题这二者导出的．为说明有约束的形式，仍设我们有(13.1)和(13.2)所定义的 m 个约束 $(m\geqslant n)$，包括显式的和隐式的，我们定义与(13.1)相联系的“边界带”如下：如果

$$l_j\leqslant x_j\leqslant l_j+\varepsilon(u_j-l_j), \tag{13.4}$$

就说分量 $x_j(j=1,\cdots,m)$ 在下边界带中．类似地，如果

$$u_j\geqslant x_j\geqslant u_j-\varepsilon(u_j-l_j), \tag{13.5}$$

则说 x_j 在上边界带中，其中 ε 是一个小常数．对于下边界带，我们定义侵入的相对宽度为

$$\eta_j=\frac{l_j+\varepsilon(u_j-l_j)-x_j}{\varepsilon(u_j-l_j)}, \tag{13.6}$$

对上边界带，类似的表达式为

$$\eta_j=\frac{x_j+\varepsilon(u_j-l_j)-u_j}{\varepsilon(u_j-l_j)},\tag{13.7}$$

以 $\hat{f}$ 表示用这方法算得的目标函数 f 在能行点（即满足(13.1)的点）上的当前最小值，以 f^+ 表示 f 在不属于边界带的能行点上的当前最小值. 显然，$f^+\geqslant\hat{f}$. 设给定一个不在任何边界带中的出发点 x_B^0，最初令 $f^+=\hat{f}=f(x_B^0)$. 象无约束形式那样进行搜索，只是在每次函数值计算之外要进行一些附加的计算. 令 $\tilde{x}$ 记按照无约束方法需要求 f 值的下一个点，则

1. 如果 $\tilde{x}$ 非能行，则称向 $\tilde{x}$ 的移动是失败，按无约束形式继续下去.

2. 如果 $\tilde{x}$ 能行且不在任何边界带中，则遵照无约束形式去做.

3. 如果 $\tilde{x}$ 在某边界带中，则按下面方式修正函数值 $f(\tilde{x})$：设 $\tilde{x}$ 落在边界带中的第一个分量的足标 j 为 j_1，按下式将 $f(\tilde{x})$ 改取为 $\bar{f}(\tilde{x})$：

$$\bar{f}(\tilde{x})=f(\tilde{x})-[f(\tilde{x})-f^+][3\eta_{j_1}-4(\eta_{j_1})^2+2(\eta_{j_1})^3],\tag{13.8}$$

增加 j 使之等于 $\tilde{x}$ 的落在边界带中的第二个分量的足标，在(13.8)的右边，利用由以前的 j 得到的 $\bar{f}(\tilde{x})$ 重新计算 $\bar{f}$. 对所有这样的 j 重复进行.

对情形(3)提出的程序是基于这样的假定：如果无约束算法导致搜索进入边界区，那末这函数的无约束极小值点可能在一个非能行点上达到. 为了使搜索限制在边界带，就在一个能行点上构造一个人为的极小值点，这构造利用了上面指出的函数修正. 假定沿着一条从能行一侧进入边界带的某直线，目标函数值是递减的. 从(13.8)可看到，对 $\eta=0$，有 $\bar{f}(\tilde{x})=f(\tilde{x})$，对 $\eta=1$，即在能行集的边界上，$\bar{f}(\tilde{x})=f^+\geqslant f(\tilde{x})$. 这样，函数值被人为地抬高以产生所希望的效果. 注意，Rosenbrock 方法的这一背景极相似于前章讨论的内部惩罚函数法的基本思想.

Davies、Swann 和 Campey 对于旋转方向法的修改也被扩充到有约束问题. 但是, 类似且更有效的方法将在下一节中给出, 所以在这里不讨论这方法了, 有兴趣的读者可参看 Davies[6] 和 Davies、Swann[7].

总之,无约束极小化的经验方法对有约束问题的扩充,是十分可靠但比较缓慢的方法,因此只在导数的计算不合用时,才值得推荐. 在近年的文献中, 关于有约束的经验方法很少出现新思想. 要使这种方法可以与惩罚函数法或下面将介绍的某些方法相竞争,还有很多工作要做.

13.2 线性约束的梯度投影算法

最速下降法是无约束极小化的基本的数值方法之一, 但我们已经看到,它不是最有效的. Rosen[33, 34] 把它扩充到约束问题,可认为这是建立**梯度投影算法**这类方法过程中的基本工作之一. 读者从第 10 章与第 11 章已经清楚,对于无约束极小化,共轭梯度算法或变尺度算法要比最速下降法远为有效. 所以, 代替收敛得慢的最速下降法,我们将讲述 Goldfarb[14] 的更近期的工作(也见 Goldfarb、Lapidus[16]), 它类似于 Rosen 的工作,但优点是它能作为共轭梯度法和变尺度法对线性约束的扩充.

考虑线性约束的非线性规划问题

(LCP) $$\min f(x) \tag{13.9}$$

受限制于

$$\sum_{j=1}^{n} a_{ij}x_j - b_i = 0, \quad i=1, \cdots, m, \tag{13.10}$$

$$\sum_{j=1}^{n} a_{ij}x_j - b_i \geqslant 0, \quad i=m+1, \cdots, p. \tag{13.11}$$

这里, 令 X 是满足(13.10)和(13.11)的点 x 之集, 并假定 f 是定义在包含 X 的 R^n 中子集 S 上的连续可微实函数. 数 a_{ij} 和 b_i 是给定的,且假定

$$\sum_{j=1}^{n} (a_{ij})^2 = 1, \quad i=1, \cdots, p, \tag{13.12}$$

令

$$a^i=(a_{i1},\ \cdots,\ a_{in})^T,\quad i=1,\ \cdots,\ p. \tag{13.13}$$

则从(13.12)推得

$$(a^i)^Ta^i=1,\quad i=1,\ \cdots,\ p, \tag{13.14}$$

即向量 a^i 有单位长度.

用 L_i 记由第 i 个约束定义的超平面, 即

$$L_i=\{x:x\in R^n,\ (a^i)^Tx=b_i\},\quad i=1,\ \cdots,\ p, \tag{13.15}$$

并用 L_i^c 记由不等式约束定义的闭半空间, 即

$$L_i^c=\{x:x\in R^n,\ (a^i)^Tx\geqslant b_i\},\quad i=m+1,\ \cdots,\ p, \tag{13.16}$$

则能行集 X 由下式给出:

$$X=(\bigcap_{i=1}^{m}L_i)\cap(\bigcap_{i=m+1}^{p}L_i^c). \tag{13.17}$$

注意, 向量 a^i 正交于超平面 L_i, 且对 $i=m+1,\ \cdots,\ p$, 它们是指入半空间 L_i^c, 即指入能行集.

一组超平面称为线性独立的, 如果对应的向量 a^i (也称为法向量)是线性独立的. q 个线性独立的超平面之交是 R^n 中一个仿射集, 记为 F^q. 平行于 F^q 的 $n-q$ 维子空间(即包含原点的仿射集)记为 M^q, 如果 F^q 本身是子空间, 则显然 $F^q=M^q$. 令 Q 是一个 n 阶正定阵. 由向量 $Q^{-1}a^i$ 张成的 q 维子空间记为 $\hat{M}^q$, 这里 a^i 对应于 F^q. 因为 R^n 可以表示为 M^q 和 $\hat{M}^q$ 的直和, 所以每个向量 $x\in R^n$ 可写成 $x=u+v$, 其中 $u\in M^q$, $v\in\hat{M}^q$, 见 [24].

我们来求向量 $x\in R^n$ 到 M^q 和 $\hat{M}^q$ 的投影映射. 用 A_q 记 $n\times q$ 阵, 其列是 q 个线性独立向量 a^i, 这里 i 属于指标集 I_q, I_q 是 $I_p=\{1,\ 2,\ \cdots,\ p\}$ 的一个子集. 不失一般性, 假定

$$I_q=\{1,\ 2,\ \cdots,\ q\}. \tag{13.18}$$

矩阵 A_q 有秩 q, 所以 $A_q^TQ^{-1}A_q$ 有逆阵. 令 $\alpha=(\alpha_1,\ \cdots,\ \alpha_q)^T$, 向量 $v\in\hat{M}^q$ 可表示为

$$v=Q^{-1}A_q\alpha, \tag{13.19}$$

而对任意 $x\in R^n$, 因为 $A_q^Tu=0$, 得

$$A_q^Tx=A_q^Tu+A_q^Tv=A_q^TQ^{-1}A_q\alpha, \tag{13.20}$$

因此有

$$\alpha=(A_q^T Q^{-1} A_q)^{-1} A_q^T x, \tag{13.21}$$

$$v=Q^{-1} A_q (A_q^T Q^{-1} A_q)^{-1} A_q^T x. \tag{13.22}$$

定义 n 阶方阵 $\hat{P}_q$ 为

$$\hat{P}_q=Q^{-1} A_q (A_q^T Q^{-1} A_q)^{-1} A_q^T, \tag{13.23}$$

则对任意 $v \in \hat{M}^q$, 有

$$\hat{P}_q v=Q^{-1} A_q (A_q^T Q^{-1} A_q)^{-1} A_q^T Q^{-1} A_q \alpha \tag{13.24}$$

$$=Q^{-1} A_q \alpha=v, \tag{13.25}$$

而对任意 $u \in M^q$, 有

$$\hat{P}_q u=Q^{-1} A_q (A_q^T Q^{-1} A_q)^{-1} A_q^T u=0. \tag{13.26}$$

这样, $\hat{P}_q$ 是将 R^n 映射到 $\hat{M}^q$ 上的**投影矩阵**[24]. 类似地, 由式

$$P_q=I-\hat{P}_q \tag{13.27}$$

定义的相补投影矩阵 P_q 把 R^n 投影到子空间 M^q 上. 在 $Q=I$ 的特殊情况, 子空间 M^q 和 $\hat{M}^q$ 互为正交补, 于是 P_q 和 $\hat{P}_q$ 是对称正交投影矩阵. 这样,

$$\hat{P}_q=A_q(A_q^T A_q)^{-1} A_q^T=(A_q^+)^T A_q^T, \tag{13.28}$$

其中 A_q^+ 是 A_q 的广义逆.

例 13.2.1

为说明前面的概念, 令 $Q=I$, 且

$$F^1=\left\{x: x \in R^2,\ -\frac{3}{5} x_1-\frac{4}{5} x_2+\frac{8}{5}=0\right\}, \tag{13.29}$$

则

$$M^1=\left\{x: x \in R^2,\ -\frac{3}{5} x_1-\frac{4}{5} x_2=0\right\}, \tag{13.30}$$

$$\hat{M}^1=\left\{x: x \in R^2,\ x_2=\frac{4}{3} x_1\right\}, \tag{13.31}$$

我们有 $A_1^T=\left(-\frac{3}{5},\ -\frac{4}{5}\right)$和

$$\hat{P}_1=A_1(A_1^T A_1)^{-1} A_1^T=\begin{pmatrix}-\frac{3}{5}\\ -\frac{4}{5}\end{pmatrix}\left(-\frac{3}{5},\ -\frac{4}{5}\right) \tag{13.32}$$

$$=\frac{1}{25}\begin{bmatrix}9 & 12\\ 12 & 16\end{bmatrix}. \tag{13.33}$$

所以

$$P_1 = I - \hat{P}_1 = \frac{1}{25}\begin{bmatrix} 16 & -12 \\ -12 & 9 \end{bmatrix}. \tag{13.34}$$

现在取 $x=(2,\ 1)^T$，则

$$\hat{P}_1 x = \left(\frac{6}{5},\ \frac{8}{5}\right)^T \in \hat{M}^1,$$

$$P_1 x = \left(\frac{4}{5},\ -\frac{3}{5}\right)^T \in M^1,$$

见图 13.1. ▌

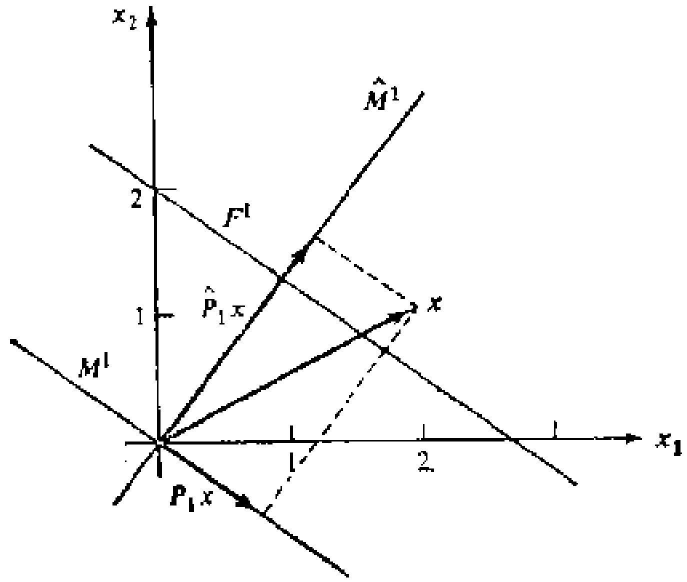


图 13.1 到正交子空间中的投影

在计算过程中，必须对不同的子空间求得投影矩阵，这些子空间平行于超平面的交集且对应于某个当前的积极不等式约束集（一个不等式约束 $g(x)\geqslant 0$ 称为在 $\hat{x}$ 是积极的，如果 $g(\hat{x})=0$）。每个新的子空间同原先子空间的差异或者是从交集中丢掉一个积极约束，或者是增加一个新的积极约束。对于这些计算，重要的是建立一个递推关系式，使得 $(A_q^T A_q)^{-1}$ 可以容易地从 $(A_{q-1}^T A_{q-1})^{-1}$ 算得，并且反过来也是一样。设 q 阶非异方阵 $(A_q^T A_q)^{-1}$ 分块为

$$(A_q^T A_q)^{-1} = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}, \tag{13.35}$$

其中 B_1、B_2、B_3、B_4 分别是 $(q-1)\times(q-1)$、$(q-1)\times 1$、$1\times(q-1)$、

1×1 矩阵. 设 a^q 是第 q 个(最后一个)向量,它相应于超平面 L_q,当 a^q 从当前的矩阵 A_q 中去掉时, 我们可以按照下式计算 $q-1$ 阶方阵 $(A_{q-1}^TA_{q-1})^{-1}$:

$$(A_{q-1}^TA_{q-1})^{-1}=B_1-B_2B_4^{-1}B_3. \tag{13.36}$$

如果不是 L_q, 而是超平面 $L_h(h<q)$ 被去掉, 同样的公式可以利用, 只要在应用 (13.36) 之前交换 $(A_q^TA_q)^{-1}$ 的第 h 行与列和第 q 行与列.

增加相应于超平面 L_q 的向量 a^q, 从 $(A_{q-1}^TA_{q-1})^{-1}$ 计算 $(A_q^TA_q)^{-1}$ 这一相反的过程如下所述: 计算

$$r_{q-1}=(A_{q-1}^TA_{q-1})^{-1}(A_{q-1}^Ta^q), \tag{13.37}$$

$$P_{q-1}a^q=a^q-A_{q-1}r_{q-1}. \tag{13.38}$$

则有

$$A_0=(P_{q-1}a^q)^T(P_{q-1}a^q), \tag{13.39}$$

$$B_1=(A_{q-1}^TA_{q-1})^{-1}+A_0^{-1}r_{q-1}r_{q-1}^T, \tag{13.40}$$

$$B_2=B_3^T=-A_0^{-1}r_{q-1}, \tag{13.41}$$

$$B_4=A_0^{-1}>0, \tag{13.42}$$

其中 $B_4>0$ 是由于 $P_{q-1}a^q\neq0$. 矩阵$(A_q^TA_q)^{-1}$ 就由(13.35)给出. 另一有用的关系式是投影矩阵的递推公式

$$P_q=P_{q-1}-\frac{P_{q-1}a^q(a^q)^TP_{q-1}}{(a^q)^TP_{q-1}a^q}. \tag{13.43}$$

此外, 可以证明, 对任意向量 $x\in R^n$,

$$\|P_qx\|\leqslant\|P_{q-1}x\|\leqslant\|x\|, \tag{13.44}$$

即投影不增加向量的长度.

Bartels、Golub、Saunders [1] 提出修改矩阵 $(A_q^TA_q)^{-1}$ 的另一过程, 它利用 Householder 变换[17]. 这个结果在数值上是更稳定的过程, 但有时要花费更长的计算.

例 13.2.2

设有两个 R^3 中的等式约束

$$\frac{2}{3}x_1-\frac{2}{3}x_2-\frac{1}{3}x_3=-\frac{1}{3}, \tag{13.45}$$

$$\frac{2}{3}x_1+\frac{1}{3}x_2+\frac{2}{3}x_3=\frac{2}{3}, \tag{13.46}$$

我们希望计算投影矩阵 P_1 和 P_2. 从(13.43)得到

$$P_1=P_0-\frac{P_0a^1(a^1)^TP_0}{(a^1)^TP_0a^1}, \tag{13.47}$$

现在令 $P_0=I$, $a^1=\left(\frac{2}{3},\ -\frac{2}{3},\ -\frac{1}{3}\right)^T$, 所以

$$P_1=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}-\frac{1}{9}\begin{bmatrix}4&-4&-2\\-4&4&2\\-2&2&1\end{bmatrix}=\frac{1}{9}\begin{bmatrix}5&4&2\\4&5&-2\\2&-2&8\end{bmatrix}, \tag{13.48}$$

类似地, $a^2=\left(\frac{2}{3},\ \frac{1}{3},\ \frac{2}{3}\right)^T$ 以及

$$P_2=P_1-\frac{P_1a^2(a^2)^TP_1}{(a^2)^TP_1a^2}, \tag{13.49}$$

将数值代入(13.49), 我们得到

$$P_2=\frac{1}{9}\begin{bmatrix}5&4&2\\4&5&-2\\2&-2&8\end{bmatrix}-\frac{1}{9}\begin{bmatrix}4&2&4\\2&1&2\\4&2&4\end{bmatrix}=\frac{1}{9}\begin{bmatrix}1&2&-2\\2&4&-4\\-2&-4&4\end{bmatrix}. \tag{13.50}$$

为了校核这些投影矩阵, 取 R^3 中任意向量, 比如 $x=(1,\ 1,\ 1)^T$, 则

$$P_1x=\begin{pmatrix}\frac{11}{9}\\ \frac{7}{9}\\ \frac{8}{9}\end{pmatrix} \quad 和 \quad P_2x=\begin{pmatrix}\frac{1}{9}\\ \frac{2}{9}\\ -\frac{2}{9}\end{pmatrix}. \tag{13.51}$$

读者容易验证 $P_1x\in M^1$ 和 $P_2x\in M^2$, 其中

$$M^1=\left\{x:x\in R^3,\ \frac{2}{3}x_1-\frac{2}{3}x_2-\frac{1}{3}x_3=0\right\}, \tag{13.52}$$

$$M^2=M^1\cap\left\{x:x\in R^3,\ \frac{2}{3}x_1+\frac{1}{3}x_2+\frac{2}{3}x_3=0\right\}, \tag{13.53}$$

还有 $\|x\|=\sqrt{3}$，$\|P_1x\|=\sqrt{\frac{26}{9}}$ 和 $\|P_2x\|=\frac{1}{3}$. 这样,(13.44)成立. ∎

也可以证明,线性相关的超平面容易借助递推公式(13.43)来发现,例如,如果发现对某 l, $P_{l-1}(a^l)=0$,则相应于 a^l 的约束就可以从问题(LCP)中去掉而不影响其解,因为它线性相关于某些其他约束.

现在介绍 Goldfarb 共轭梯度投影算法. 这方法将给出而不予证明. 关于进一步的细节和证明,有兴趣的读者可参考[14]和[16].

设给定一个初始能行点 $x^0\in X$. 显然,如果出现等式约束,那末有 $x^0\in L_i$, $i=1, \cdots, m$. 所以,对某 $q\geqslant m$, $x^0\in F^q$, 不失一般性可假定 $A_q=(a^1, \cdots, a^m, \cdots, a^q)$. 并设给定了一个 n 阶正定对称阵 H_0^0. 用下面递推关系式计算 H_q^0:

$$H_i^0=H_{i-1}^0-\frac{H_{i-1}^0a^i(a^i)^TH_{i-1}^0}{(a^i)^TH_{i-1}^0a^i},\quad i=1, \cdots, q, \tag{13.54}$$

其中相应于等式约束(13.10)的 $a^i(i=1, \cdots, m)$ 首先加入. 令 $k=0, 1, \cdots$, 在第 k 次迭代中,执行如下步骤:

1. 给定一个能行点 x^k, $\nabla f(x^k)$ 和 H_q^k, 计算 $H_q^k\nabla f(x^k)$. 对 $q\geqslant 1$ 计算

$$\lambda^k=(A_q^TA_q)^{-1}A_q^T\nabla f(x^k), \tag{13.55}$$

若 x^k 不在相应于第 i 个不等式约束的超平面 L_i 中, 则令 $\lambda_i^k=0$. 若

$$H_q^k\nabla f(x^k)=0 \tag{13.56}$$

且

$$\lambda_i^k\geqslant 0,\quad i=m+1, \cdots, p, \tag{13.57}$$

则 x^k 满足(LCP)中最优性的一阶必要条件,算法就结束.

2. 若(13.56)和(13.57)不同时满足,则或者

$$\|H_q^k\nabla f(x^k)\|>\max\left\{0,\ -\frac{1}{2}\lambda_q^kb_{qq}^{-1/2}\right\}, \tag{13.58}$$

或者

$$\|H_q^k\nabla f(x^k)\| \leqslant -\frac{1}{2}\lambda_q^k b_{qq}^{-1/2}, \quad \lambda_q^k<0, \tag{13.59}$$

其中 b_{ii} 是 $(A_q^TA_q)^{-1}$ 的第 i 个对角元素，为方便起见，假定

$$\lambda_q^k b_{qq}^{-1/2} \leqslant \lambda_q^k b_{ii}^{-1/2}, \quad i=m+1,\ \cdots,\ q, \tag{13.60}$$

注意每个 $b_{ii}>0$. 若 (13.58) 成立，则进入步骤 3, 否则从积极约束集中去掉第 q 个约束，如早先所述，从 $(A_q^TA_q)^{-1}$ 计算 $(A_{q-1}^TA_{q-1})^{-1}$, 并得到

$$P_{q-1}=I-A_{q-1}(A_{q-1}^TA_{q-1})^{-1}A_{q-1}^T. \tag{13.61}$$

然后，计算

$$H_{q-1}^k=H_q^k-\frac{P_{q-1}a^q(a^q)^TP_{q-1}}{(a^q)^TP_{q-1}a^q}. \tag{13.62}$$

令 $q=q-1$, 回到步骤 1.

3．令 $z^{k+1}=-H_q^k\nabla f(x^k)$ (注意，在第 11 章中我们用 $\nabla f(x^k)$ 乘以 H 的转置来定义 z^{k+1}, 为了简化记号，在描述这个算法时省去了转置号). 计算 β_i^{k+1}, 它是沿着 z^{k+1} 而不违反第 i 个当前的非积极约束的最大步长，

$$\beta_i^{k+1}=-\frac{(a^i)^Tx^k-b^i}{(a^i)^Tz^{k+1}}, \quad i=q+1,\ \cdots,\ p, \tag{13.63}$$

所以，沿着 z^{k+1} 在能行域中能取的最大步长为

$$\beta_{\max}^{k+1}=\min_i\{\beta_i^{k+1}:\beta_i^{k+1}>0,\ i=q+1,\ \cdots,\ p\}. \tag{13.64}$$

进行一个线性搜索求 α_{k+1}^*, $0\leqslant\alpha_{k+1}^*\leqslant\beta_{\max}^{k+1}$, 它使

$$f(x^k+\alpha_{k+1}z^{k+1})$$

取极小. 有可能 $\alpha_{k+1}^*\to+\infty$, 这就表示 f 在 z^{k+1} 方向上是无界的. 如果 $\alpha_{k+1}^*<+\infty$, 令

$$x^{k+1}=x^k+\alpha_{k+1}^*z^{k+1} \tag{13.65}$$

且计算 $\nabla f(x^{k+1})$.

4．如果 $\alpha_{k+1}^*<\beta_{\max}^{k+1}$, 则进入步骤 5, 否则把相应于 $\beta_{\max}^{k+1}$ 的约束(比如是第$(q+1)$个约束)加入积极(约束)集，并计算

$$H_{q+1}^{k+1}=H_q^k-\frac{H_q^ka^{q+1}(a^{q+1})^TH_q^k}{(a^{q+1})^TH_q^ka^{q+1}}. \tag{13.66}$$

令 $q=q+1$ 和 $k=k+1$, 回到步骤 1.

5. 在现在的情形，积极约束集保持不变，用第 11 章的秩 1 或秩 2 变尺度公式，比如 BFS 公式，修改 H_q^k:

$$H_q^{k+1}=H_q^k+\left[1+\frac{(\gamma^{k+1})^T H_q^k \gamma^{k+1}}{(p^{k+1})^T \gamma^{k+1}}\right]\frac{p^{k+1}(p^{k+1})^T}{(p^{k+1})^T\gamma^{k+1}}-\frac{p^{k+1}(\gamma^{k+1})^T H_q^k}{(p^{k+1})^T\gamma^{k+1}}-\frac{H_q^k\gamma^{k+1}(p^{k+1})^T}{(p^{k+1})^T\gamma^{k+1}}, \tag{13.67}$$

如以前一样，其中

$$p^{k+1}=x^{k+1}-x^k,\quad \gamma^{k+1}=\nabla f(x^{k+1})-\nabla f(x^k),$$

令 $k=k+1$，回到步骤 1.

这就完成了算法的描述.

例 13.2.3

考虑问题

$$\min f(x)=(x_1)^2+4(x_2)^2 \tag{13.68}$$

受限制于

$$\frac{3}{5}x_1+\frac{4}{5}x_2\geqslant\frac{13}{5}. \tag{13.69}$$

我们用求解这个二次规划来说明 Goldfarb 算法. 设 $x^0=(0,5)^T$，其函数值 $f(x^0)=100$，我们选取

$$H_0^0=\begin{bmatrix}1&0\\0&1\end{bmatrix}. \tag{13.70}$$

目标函数的梯度为

$$\nabla f(x)=\begin{pmatrix}2x\\8x\end{pmatrix}, \tag{13.71}$$

在 x^0 处有

$$\nabla f(x^0)=\begin{pmatrix}0\\40\end{pmatrix}. \tag{13.72}$$

点 x^0 以严格不等式满足(13.69)，所以 $\lambda_1^0=0$ 且 $\|H_0^0\nabla f(x^0)\|=40$. 我们进入步骤 3，得到

$$z^1=-H_0^0\nabla f(x^0)=\begin{pmatrix}0\\-40\end{pmatrix}, \tag{13.73}$$

并计算在这方向上不违反约束(13.69)所能移动的最大距离.

$$\beta_1^1 = -\frac{(3/5)(0)+(4/5)(5)-(13/5)}{(3/5)(0)+(4/5)(-40)} = \frac{7}{160}. \tag{13.74}$$

读者容易验证，沿着 z^1 方向在区间 $0 \leqslant \alpha_1 \leqslant \frac{7}{160}$ 中进行线性搜索，可得 $\alpha_1^* = \beta_{\max}^1 = \frac{7}{160}$，并且

$$x^1 = \begin{pmatrix} 0 \\ 5 \end{pmatrix} + \frac{7}{160}\begin{pmatrix} 0 \\ -40 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{13}{4} \end{pmatrix}, \tag{13.75}$$

$f(x^1) = \frac{168}{4}$. 约束(13.69)现在变成积极的，由(13.66)计算

$$H_1^1 = H_0^0 - \frac{H_0^0 a^1 (a^1)^T H_0^0}{(a^1)^T H_0^0 a^1} \tag{13.76}$$

$$= \frac{1}{25}\begin{bmatrix} 16 & -12 \\ -12 & 9 \end{bmatrix}. \tag{13.77}$$

又令 $q=1$, $k=1$，且回到步骤1以作最优性的检验. 我们有

$$(A_1^T A_1)^{-1} = 1, \quad \nabla f(x^1) = (0,\ 26)^T, \quad \lambda_1^1 = \frac{104}{5}.$$

现在

$$H_1^1 \nabla f(x^1) = \begin{pmatrix} -\frac{312}{25} \\ \frac{234}{25} \end{pmatrix}. \tag{13.78}$$

这样，x^1 不是最优解. 因为 $\| H_1^1 \nabla f(x^1) \| = \frac{78}{5}$ 且(13.58)满足，下一个搜索方向就和负梯度在相应于约束的超平面上的投影相重合，即

$$z^2 = -H_1^1 \nabla f(x^1) = \begin{pmatrix} \frac{312}{25} \\ -\frac{234}{25} \end{pmatrix}. \tag{13.79}$$

因为没有其他约束，我们有 $\beta_{\max}^2 = +\infty$，进行一个线性搜索可求得 $\alpha_2^* = \frac{25}{104}$，所以

$$x^2=\begin{pmatrix}0\\ \dfrac{13}{4}\end{pmatrix}+\frac{25}{104}\begin{pmatrix}\dfrac{312}{25}\\ -\dfrac{234}{25}\end{pmatrix}=\begin{pmatrix}3\\ 1\end{pmatrix}, \tag{13.80}$$

$$\nabla f(x^2)=\begin{pmatrix}6\\ 8\end{pmatrix}. \tag{13.81}$$

函数值 $f(x^2)=13$，可以看到这方法是下降的．现在用(13.67)来修改 H_1^1．首先计算

$$p^2=\begin{pmatrix}3\\ 1\end{pmatrix}-\begin{pmatrix}0\\ \dfrac{13}{4}\end{pmatrix}=\begin{pmatrix}3\\ -\dfrac{9}{4}\end{pmatrix}, \tag{13.82}$$

$$\gamma^2=\begin{pmatrix}6\\ 8\end{pmatrix}-\begin{pmatrix}0\\ 26\end{pmatrix}=\begin{pmatrix}6\\ -18\end{pmatrix}. \tag{13.83}$$

将(13.77)，(13.82)和(13.83)代入(13.67)，产生

$$H_1^2=\frac{1}{104}\begin{bmatrix}16 & -12\\ -12 & 9\end{bmatrix}, \tag{13.84}$$

结果是

$$H_1^2\nabla f(x^2)=\frac{1}{104}\begin{bmatrix}16 & -12\\ -12 & 9\end{bmatrix}\begin{pmatrix}6\\ 8\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix}, \tag{13.85}$$

且

$$\lambda_1=\left(\frac{3}{5}\right)(6)+\left(\frac{4}{5}\right)(8)=10. \tag{13.86}$$

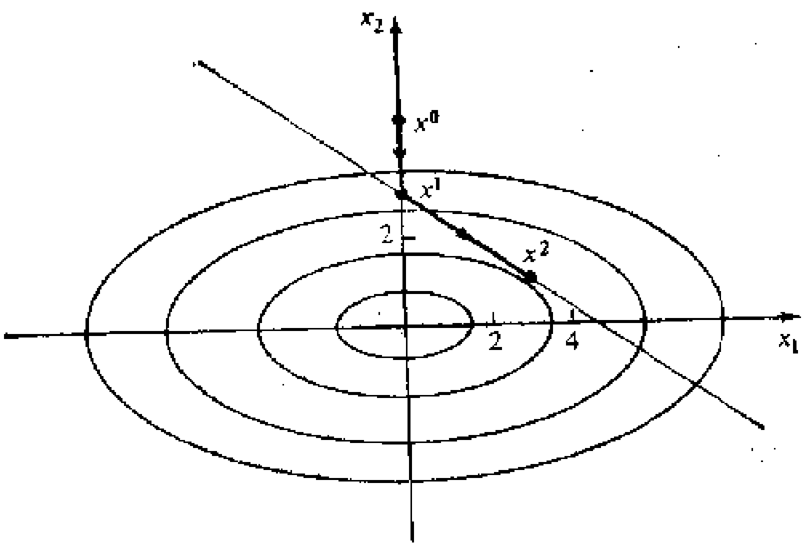


图 **13.2** Goldfarb 的算法

点 $x^2=(3,1)^T$ 满足(13.56)和(13.57)，所以它是最优的，最优化的行程在图 13.2 中表明. ▌

Goldfarb[14] 证明，前面的算法一般是一个下降方法，换言之，在每次迭代时，目标函数值是减少的. 最优性条件也可用矩阵 H_q 来表述. 这条件在下述定理中给出.

定理 13.1

设 f 是一个连续可微实函数，$x^*\in X$ 是位于 F^q 中的 X 的边界点，F^q 是由约束(13.10)和(13.11)所定义的 q 个线性独立超平面之交. 如果 x^* 是问题(LCP)的局部最优点，那末

$$H_q\nabla f(x^*)=0, \tag{13.87}$$

$$\lambda_i^*\geqslant 0,\quad i=m+1,\ \cdots,\ q, \tag{13.88}$$

其中

$$\lambda^*=(\lambda_1^*,\ \cdots,\ \lambda_q^*)^T=(A_q^TA_q)^{-1}A_q^T\nabla f(x^*). \tag{13.89}$$

如果 f 还是凸的，那末(13.87)至(13.89)是关于整体约束极小值点的充分条件.

阅读这定理的证明可以帮助读者理解上面给出的算法步骤. 关于证明以及梯度投影法的很多其他细节，见 Rosen [33] 和 Goldfarb [14]，这里不能涉及了.

使用 Goldfarb 方法的数值经验是良好的，但对于这算法的上述形式，即使在凸性假定下，也不能在理论上保证收敛到(LCP)的最优解. Rosen 的梯度投影法的某种更复杂的形式由 Polak[25, 26] 提出，他证明了在 f 是凸的且满足某些正则性条件时，收敛到最优解.

分析(LCP)中目标函数 f 是严格凸的二次函数且只有等式约束(13.10)的情形，可以对梯度投影法得到一个有趣的理解. 如 Fletcher[10] 那样，令

$$f(x)=a+c^Tx+\frac{1}{2}x^TQx. \tag{13.90}$$

令 A_m 是相应于约束(13.10)的 $n\times m$ 矩阵. 这样一个二次规划的最优解 x^* 和相应的 Lagrange 乘子 λ^* 由下列线性方程组的解给出:

$$\begin{bmatrix} Q & -A_m \\ -A_m^T & 0 \end{bmatrix}\begin{pmatrix} x \\ \lambda \end{pmatrix}+\begin{pmatrix} c \\ b \end{pmatrix}=0. \tag{13.91}$$

注意,

$$\begin{bmatrix} Q & -A_m \\ -A_m^T & 0 \end{bmatrix}^{-1}$$

$$=\begin{bmatrix} Q^{-1}-Q^{-1}A_m(A_m^TQ^{-1}A_m)^{-1}A_m^TQ^{-1} & -Q^{-1}A_m(A_m^TQ^{-1}A_m)^{-1} \\ -(A_m^TQ^{-1}A_m)^{-1}A_m^TQ^{-1} & -(A_m^TQ^{-1}A_m)^{-1} \end{bmatrix} \tag{13.92}$$

$$=\begin{bmatrix} H_m^* & -A_m^{*T} \\ -A_m^* & -(A_m^TQ^{-1}A_m)^{-1} \end{bmatrix}, \tag{13.93}$$

其中 H_m^* 和 A_m^* 只是一个缩写,我们可以把(13.91)写成

$$x^*=x^0-H_m^*\nabla f(x^0), \tag{13.94}$$

$$\lambda^*=A_m^*\nabla f(x^0), \tag{13.95}$$

其中 x^0 是任意能行点.

我们也可以写

$$H_m^*=[I-Q^{-1}A_m(A_m^TQ^{-1}A_m)^{-1}A_m^T]Q^{-1}, \tag{13.96}$$

并且,利用(13.23)和(13.27)我们有

$$H_m^*=P_mQ^{-1}. \tag{13.97}$$

所以 H_m^* 可以看作是 f 的投影 Hesse 逆阵. 令 $Q=I$, 我们得到 $H_m^*=P_m$, 这样重新得到 Rosen 的梯度投影法.

在一般非线性函数的情形, (13.94)不是精确的, 所以要求一个迭代过程以求得最优值点 x^*, 例如用迭代

$$x^{k+1}=x^k-\alpha_{k+1}^*H_k^*\nabla f(x^k), \tag{13.98}$$

其中 H_k^* 是近似投影 Hesse 逆阵, 象 Goldfarb 方法那样每次迭代时要修改之.

基于(13.96)和(13.98)的梯度投影法由 Murtagh、Sargent[22] 导出,它利用搜索方向

$$\begin{aligned} z^{k+1} &= -H_k^{*T}\nabla f(x^k) \\ &= -[H_k^T-H_k^TA_m(A_m^TH_k^TA_m)^{-1}A_m^TH_k]\nabla f(x^k), \end{aligned} \tag{13.99}$$

其中 H_k^T 是 Hesse 逆阵的当前近似. 这矩阵的修改按秩 1 校正矩阵来进行(见第 11 章). 类似于 Goldfarb 方法, Murtagh-Sargent 算法也利用一个"积极约束集策略"——在迭代中, 当从一个能行点移动到另一能行点时, 如必要的话增加或丢掉约束. Goldfarb 和 Murtagh-Sargent 方法都具有考虑到约束而适当修正了的二次终结性质. 对于 Goldfarb 方法, 我们有下面的结果, 它是属于 Powell 的[27, 28].

定理 13.2

如果利用 Goldfarb 方法来求解一个具有 n 阶正定阵 Q 的二次函数 f 在线性约束(13.10)和(13.11)下的极小化问题, 并且如果在计算过程中积极约束集改变了 l 次, 那末 f 的精确极小值点可以通过不超过 $n+l$ 次精确的线性搜索求得.

Powell[29] 证明, Murtagh-Sargent 方法可看作是广义 Goldfarb 算法的一种特殊情形且反之亦然, 从而推广了这定理. 设(13.99)中的矩阵 H_k 由下式修改:

$$H_k = H_k^{\mathrm{BFS}} + \beta^k w^k (w^k)^T, \tag{13.100}$$

其中 H_k^{BFS} 如(11.170)给出,是利用 Broyden-Fletcher-Shanno 的拟 Newton 法的修改公式得到的, w^k 由 (11.169) 定义. 参数 β^k 是以互补的方式联系于(11.167)中出现的 α_{12}^k, 这是因为对于 BFS 方法 $\beta^k=0$, 而对于 DFP 方法 $\alpha_{12}^k=0$. 其次设 H_k 由 (13.100) 给出, 并令

$$H_q^k = H_k - H_k A_q (A_q^T H_k A_q)^{-1} A_q^T H_k. \tag{13.101}$$

因为 $H_q^k A_q = 0$, 矩阵 H_q^k 就是 Goldfarb 方法的一个合适的投影矩阵. 再设用 Murtagh-Sargent 类型的方法求解一个具有正定阵 Q 的二次函数的极小化问题, 利用具有任意 β^0, $x^0 \in R^n$, 和非异阵 H_0 的一般修改公式(13.100), 参数 β^k 和 α_k (步长)在每次迭代时任意取. 现在用 Goldfarb 方法, 从同样的 $x^0 \in R^n$ 出发, H_q^0 由(13.101)给出, 其中 α_k 在每次迭代时用同样的值, 并且

$$H_q^k = H_q^{(\mathrm{BFS})k} + \hat{\beta}^k w^k (w^k)^T, \tag{13.102}$$

其中 $\hat{\beta}^k$ 对 β^k 的关系为

$$\hat{\beta}^k = \frac{\beta^k}{1+\beta^k(\gamma^k)^T H_k A_q (A_q^T H_k A_q)^{-1} A_q^T H_k \gamma^k},$$

(13.103)

$H_q^{(BFS)k}$ 由 BFS 方法的修改公式得到. 可以证明，这两个方法产生同样的点列. 增加或丢掉约束并不影响产生同样的点，因此，Murtagh-Sargent 方法和有关方法有定理 13.2 所述的二次终结性质.

13.3 一个二次规划算法

下一步讨论求解线性约束问题的投影梯度法的一种特殊情形. 设要极小化的目标函数是一个严格凸的二次函数，约束由线性等式和不等式给出. 为了简单起见，我们仅考虑不等式约束，因为它们在积极约束集策略中是重要的. 如果等式约束也出现，读者容易修改有关结果. 叙述方式将按照 Fletcher[9] 和 Goldfarb[14] 的著作，前者更一般些，它也能用于非凸二次函数. 关于更一般方法的详细叙述以及和其他求解二次规划的现有方法的比较，读者可参考[9]. 二次规划是在第 7 章提到的，除了在应用方面有用之外，它还作为这里和下一章中叙述的相当一些算法中的子问题.

考虑问题

(SCQP) $$\min f(x) = a + c^T x + \frac{1}{2} x^T Q x \tag{13.104}$$

受限制于

$$A^T x \geqslant b, \tag{13.105}$$

其中 $c \in R^n$, $x \in R^n$, Q 是实的对称正定阵，A 是 $n \times m$ 实矩阵，$b \in R^m$，我们将通过产生一个二次规划的序列来求解 (SCQP)，这些二次规划是受那些代表积极约束的等式所约束的. 这序列的组成者的差别仅在于加到积极约束集或从积极约束集中丢掉的那些约束. 这样的一个等式约束问题写成

(ECQP) $$\min f(x) = a + c^T x + \frac{1}{2} x^T Q x \tag{13.106}$$

受限制于

$$A_q^T x = b^q, \tag{13.107}$$

其中为简单计，$n\times q$ 阵 A_q 是由取 A 的前 q 列得到的，$b^q=(b_1,\cdots,b_q)^T$. 我们假定 A_q 的秩是 q. 这问题的解和相应的 Lagrange 乘子由(13.94)和(13.95)明显地给出，即

$$x^* = x^0 - H_q^* \nabla f(x^0), \tag{13.108}$$

$$\lambda^* = A_q^* \nabla f(x^0), \tag{13.109}$$

其中 x^0 是满足(13.107)的任意点，且

$$H_q^* = [I - Q^{-1}A_q(A_q^T Q^{-1} A_q)^{-1} A_q^T] Q^{-1}, \tag{13.110}$$

$$A_q^* = (A_q^T Q^{-1} A_q)^{-1} A_q^T Q^{-1}. \tag{13.111}$$

这算法的效率依赖于矩阵 H_q^*、A_q^* 的递推关系式. 首先假设，我们增加第 $q+1$ 个约束到积极集中，这步骤相当于把 a^{q+1}(A 的第 $q+1$ 列)加到 A_q 中，则

$$H_{q+1}^* = H_q^* - \frac{H_q^* a^{q+1}(a^{q+1})^T H_p^{*T}}{(a^{q+1})^T H_q^* a^{q+1}}, \tag{13.112}$$

$$A_{q+1}^* = \begin{bmatrix} A_q^* \\ 0, \cdots, 0 \end{bmatrix} + \begin{pmatrix} -A_q^* a^{q+1} \\ 1 \end{pmatrix} \frac{(a^{q+1})^T H_q^{*T}}{(a^{q+1})^T H_q^{*T} a^{q+1}}. \tag{13.113}$$

然后假设要从积极集中丢掉相应于 a^q 的约束. 令 $(a^{q*})^T$ 是 A_q^* 的第 q 行. 则有关系式

$$H_{q-1}^* = H_q^* + \frac{a^{q*}(a^{q*})^T}{(a^{q*})^T Q a^{q*}}, \tag{13.114}$$

$$\begin{bmatrix} A_{q-1}^* \\ 0, \cdots, 0 \end{bmatrix} = A_q^* - \frac{A_q^* Q a^{q*}(a^{q*})^T}{(a^{q*})^T Q a^{q*}}. \tag{13.115}$$

如果任何约束，比如说第 $p(<q)$ 个约束，要被丢掉，这些公式仍可用，只要在使用(13.114)和(13.115)之前将 A_q^* 的第 p 列和第 q 列交换. 注意，在 n 个积极约束时，$H_n^*=0$ (零阵)以及 $A_n^*=A^{-1}$.

前面两类矩阵修正规则可以组合成所谓交换公式，在这公式中 A_q 的第 q 列由另一个当前不在 A_q 中的向量，比如说 a^l，来代替. 作为交换运算结果，从 H_q^* 和 A_q^* 得到的新矩阵记为 H_q^{**} 和 A_q^{**}，则

$$H_q^{**}=H_q^*+\frac{a^{q*}u^T}{y}-\frac{H_q^*a^tw^T}{y}, \qquad (13.116)$$

$$A_q^{**}=A_q^*-\frac{(A_q^*a^t-e^q)w^T}{y}-\frac{A_q^*Qa^{q*}u^T}{y}, \qquad (13.117)$$

其中

$$u=a^{q*}(a^t)^TH_q^*a^t-H_q^*a^t(a^t)^Ta^{q*}, \qquad (13.118)$$

$$w=H_q^*a^t(a^{q*})^TQa^{q*}+a^{q*}(a^t)^Ta^{q*}, \qquad (13.119)$$

$$y=((a^t)^Ta^t)^2+(a^{q*})^TQa^{q*}(a^t)^TH_q^*a^t, \qquad (13.120)$$

$$e^q=(0,\ \cdots,\ 0,\ 1)^T\in R^q. \qquad (13.121)$$

这样的交换公式在 Q 是病态或 H_q^* 接近奇异时是很有用的，这时虽然最优解是有定义的，但在计算中算法可能陷入数值困难. Fletcher[9] 证明，可以利用一个在增加、丢掉和交换约束的诸修正公式中作选择的有效的总策略，以避免数值困难，而且算法也可以扩充到包含半定或不定阵 Q 的问题.

为了开始这算法，要求一个能行点. 这样一点可以采用下一节将说明的线性规划来求得，或者采用其他方法，如 Fletcher 的过程[8]，在这过程中，n 个积极约束的交集中一点(如果存在)被定出，并且算得 $A_n^*=A^{-1}$，$H_n^*=0$. 这算法的步骤现在可叙述为:

1. 给定位于 q 个积极约束的交集中的一个能行点 x^0，计算 H_q^*，A_q^*，$H_q^*\nabla f(x^0)$，令 $k=0$.

2. 计算 $z^{k+1}=-H_q^*\nabla f(x^k)$. 若 $z^{k+1}=0$，则进到步骤 4.

3. 若 $z^{k+1}\neq 0$，则计算 $x^{k+1}=x^k+\alpha_{k+1}^*z^{k+1}$ 和 $\nabla f(x^{k+1})$，其中

$$\alpha_{k+1}^*=\min\{1,\ \alpha_{k+1}\}, \qquad (13.122)$$

$$\alpha_{k+1}=\min_i\left\{-\frac{(a^i)^Tx^k-b_i}{(a^i)^Tz^{k+1}}:(a^i)^Tz^{k+1}<0,\ i=q+1,\cdots,\ m\right\}. \qquad (13.123)$$

为简单起见，设上述极小在 $i=q+1$ 时达到. 如果 $\alpha_{k+1}^*<1$，把第 $q+1$ 个约束加到积极集中，用(13.112)和(13.113)修正 H_q^*，A_q^*，令 $k=k+1$ 和 $q=q+1$，回到步骤 2; 如果 $\alpha_{k+1}^*=1$，令 $k=k+1$，进到步骤 4.

4. 计算(近似的) Lagrange 乘子向量

$$\lambda^k = A_q^* \nabla f(x^k), \tag{13.124}$$

和它的最小元素

$$\lambda_r^k = \min_i \{\lambda_i^k : i=1, \cdots, q\}. \tag{13.125}$$

为简单起见设 $r=q$，如果 $\lambda_q^k \geqslant 0$，就停止；一个最优解达到了．如果 $\lambda_q^k < 0$，从积极集中丢掉第 q 个约束并按照(13.114)和(13.115)修正 H_q^*，A_q^*．令 $q=q-1$，回到步骤 2．

例 13.3.1

我们用前面的算法去解一个二次规划问题，它取自 Beale[2]，给出如下：

$$\min f(x) = 9 - 8x_1 - 6x_2 - 4x_3 + 2(x_1)^2 + 2(x_2)^2 + (x_3)^2 + 2x_1x_2 + 2x_1x_3 \tag{13.126}$$

受限制于

$$(a^1)^T x = x_1 \geqslant 0, \tag{13.127}$$

$$(a^2)^T x = x_2 \geqslant 0, \tag{13.128}$$

$$(a^3)^T x = x_3 \geqslant 0, \tag{13.129}$$

$$(a^4)^T x = -x_1 - x_2 - x_3 \geqslant -3. \tag{13.130}$$

设我们从 $x^0 = (0, 0, 0)^T$ 出发．这样，前面三个约束是积极的，推得 $H_3^* = 0$，$A_3^* = A_3^{-1} = I$ 和 $\nabla f(x^0) = (-8, -6, -4)^T$．我们有

$$z^1 = -H_3^* \nabla f(x^0) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \tag{13.131}$$

并进到步骤 4．

计算 Lagrange 乘子，注意到 $\lambda^0 = (-8, -6, -4)^T$ 和 $\lambda_1^0 = \min\{\lambda_i^0\}$．这样，从积极集中丢掉第一个约束．因为 $r=1$，我们在 A_3^* 中对换第 1 行和第 3 行，并得到一个新矩阵

$$A_3^* = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}. \tag{13.132}$$

用(13.114)和(13.115)，计算得

$$H_2^* = \begin{bmatrix} \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_2^* = \begin{bmatrix} -\frac{1}{2} & 0 & 1 \\ -\frac{1}{2} & 1 & 0 \end{bmatrix}, \tag{13.133}$$

并回到步骤 2. 其次，我们计算

$$z^1 = -H_2^* \nabla f(x^0) = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \tag{13.134}$$

因为 $z^1 \neq 0$，我们进到步骤 3. 因为

$$(a^4)^T z^1 < 0 \quad 和 \quad 对\ i \neq 4,\ (a^i)^T z^1 \geq 0,$$

有

$$\alpha_1 = \frac{-(a^4)^T x^0 - b_4}{(a^4)^T z^1} = \frac{3}{2} > 1, \quad \alpha_1^* = 1, \tag{13.135}$$

由此可推得 $x^1 = (2,\ 0,\ 0)^T$，$\nabla f(x^1) = (0,\ -2,\ 0)^T$，进到步骤 4. 我们计算

$$\lambda^1 = \begin{bmatrix} -\frac{1}{2} & 0 & 1 \\ -\frac{1}{2} & 1 & 0 \end{bmatrix} \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \end{pmatrix}. \tag{13.136}$$

因为 $\lambda_2^1 < 0$，从积极集中丢掉第 2 个约束，再按 (13.114) 和 (13.115)，有

$$H_1^* = \begin{bmatrix} \frac{1}{3} & -\frac{1}{6} & 0 \\ -\frac{1}{6} & \frac{1}{3} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_1^* = \left[-\frac{2}{3},\ \frac{1}{3},\ 1\right]. \tag{13.137}$$

回到步骤 2，计算

$$z^2 = -H_1^* \nabla f(x^1) = \begin{pmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ 0 \end{pmatrix}, \tag{13.138}$$

因为 $z^2 \neq 0$，继续到步骤 3. 我们发现对 $i = 1,\ 4$，$(a^i)^T z^2 < 0$ 和

$$\alpha_2=\min\{6,\ 3\}>1,\quad \alpha_2^*=1, \tag{13.139}$$

所以

$$x^2=\left(\frac{5}{3},\ \frac{2}{3},\ 0\right)^T \quad 和 \quad \nabla f(x^2)=\left(0,\ 0,\ -\frac{2}{3}\right)^T.$$

计算新的 Lagrange 乘子，我们求得

$$\lambda^2=A_1^*\nabla f(x^0)=-\frac{2}{3}<0. \tag{13.140}$$

从积极集中丢掉第 3 个约束. 修正 H_q^* 和 A_q^*，我们得到

$$H_0^*=\begin{bmatrix} 1 & -\frac{1}{2} & -1 \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -1 & \frac{1}{2} & \frac{3}{2} \end{bmatrix}, \tag{13.141}$$

A_0^* 当然是空集. 回到步骤 2，我们计算

$$z^3=-H_0^*\nabla f(x^2)=\begin{pmatrix} -\frac{2}{3} \\ \frac{1}{3} \\ 1 \end{pmatrix}, \tag{13.142}$$

在步骤 3 中，得到

$$\alpha_3^*=\frac{2}{5}.$$

所以

$$x^3=\left(\frac{7}{5},\ \frac{4}{5},\ \frac{2}{5}\right)^T \quad 和 \quad \nabla f(x^3)=\left(0,\ 0,\ -\frac{2}{5}\right)^T.$$

因为 α_3^* 相应于第 4 个约束，我们把它加到(空的)积极集中去，按照(13.112)和(13.113)得到新矩阵如下：

$$H_1^*=\frac{1}{18}\begin{bmatrix} 9 & -3 & -3 \\ -3 & 5 & -1 \\ -3 & -1 & 2 \end{bmatrix},\ A_1^*=\left[\frac{1}{3},\ -\frac{2}{9},\ -\frac{5}{9}\right]. \tag{13.143}$$

下一个搜索方向为

$$z^4=-H_1^*\nabla f(x^3)=\frac{1}{45}\begin{pmatrix}-3\\-1\\2\end{pmatrix}, \tag{13.144}$$

从(13.122)和(13.123)我们求得 $\alpha_4^*=1$，所以

$$x^4=\left(\frac{4}{3},\frac{7}{9},\frac{4}{9}\right)^T \quad 和 \quad \nabla f(x^4)=(0,\ 0,\ 0)^T,$$

即最优解已经求到. ▌

如果二次函数的矩阵 Q 不是正定的，Fletcher[9] 已经导出了特殊的步骤，以保证他的算法对于一般二次规划问题的可用性. 这方法的某些改进在[30]中也可找到.

13.4 能行方向法

我们再来考察求解一大类数学规划问题的一种十分一般的途径. 设有问题

$$\min f(x) \tag{13.145}$$

受限制于

$$x\in X\subset R^n, \tag{13.146}$$

其中 X 是连通闭集，f 是具有有界水平集的连续可微函数. 求解这问题的一个算法称为**能行方向法**，如果该算法能以下述方式来描述.

对于 $k=0,\ 1,\ \cdots$，给定一点 $x^k\in X$，找一个方向向量 z^{k+1} 使得

$$(z^{k+1})^T\nabla f(x^k)<0, \tag{13.147}$$

并存在 $\bar{\alpha}_{k+1}>0$，使对一切 $0\leqslant\alpha_{k+1}\leqslant\bar{\alpha}_{k+1}$，均有 $x^k+\alpha_{k+1}z^{k+1}\in X$，这就是说，$z^{k+1}$ 是一个**能行下降方向**，或称为"可用的"方向. 按某种准则决定 α_{k+1}^*，使得

$$x^{k+1}=x^k+\alpha_{k+1}^*z^{k+1}\in X. \tag{13.148}$$

按这方式进行，一直到满足某个停止条件，可以希望它是一个收敛条件.

能行方向法由 Zoutendijk[41] 导出，在他的一系列文章中可找

到，例如见 [42, 43, 44]．读者容易看到，前面叙述的无约束算法的大部分，以及本章中给出的处理约束问题的方法都是能行方向法的特殊情形．此外，线性规划的单纯形法和将在第 14 章讨论的某些方法也可看作是能行方向法．应当提及，因为能行方向法的这个一般性，本节所述的思想和结果是否真是无约束最优化方法的扩充，从而应属于本章，这一点是成问题的．我们把本节的能行方向法的讨论限于线性约束的非线性规划情形：

(LCP) $$\min f(x) \tag{13.149}$$

受限制于

$$\sum_{j=1}^{n} a_{ij}x_j - b_i = 0, \quad i=1, \cdots, m, \tag{13.150}$$

$$\sum_{j=1}^{n} a_{ij}x_j - b_i \geqslant 0, \quad i=m+1, \cdots, p, \tag{13.151}$$

这在上节已定义了．对于能行方向法的每次迭代，要求两个主要决策：

1．选择一个能行下降方向．

2．沿着这能行方向选择一个步长．

首先要求读者证明，给定一点 $x^k \in X$（即 x^k 满足(13.150)和(13.151)），从 x^k 出发的能行下降方向 z 必定落在 $Z^1(x^k) \cap Z^2(x^k)$ 中，$Z^1(x^k) \cap Z^2(x^k)$ 是第 3 章中在更一般意义下定义的，这里

$$Z^1(x^k) = \left\{z: \sum_{j=1}^{n} a_{ij}z_j = 0,\ i=1, \cdots, m;\ \sum_{j=1}^{n} a_{ij}z_j \geqslant 0,\ i \in I(x^k)\right\}, \tag{13.152}$$

$$Z^2(x^k) = \{z: z^T \nabla f(x^k) < 0\}, \tag{13.153}$$

其中

$$I(x^k) = \left\{i: \sum_{j=1}^{n} a_{ij}x_j^k = b_i,\ i=m+1, \cdots, p\right\}. \tag{13.154}$$

一般地，从一个给定点 $x^k \in X$ 出发有很多能行下降方向，所以通过求解下面的最优化问题，可在这些方向中选取**最速能行下降方向**，这问题是

(SFD) $$\min z^T \nabla f(x^k) \tag{13.155}$$

受限制于 $z \in Z^1(x^k)$，且

$$\|z\| \leqslant 1. \tag{13.156}$$

因为 $Z^1(x^k)$ 是一个锥，必须限制 $\|z\|$ 以获得一个有限解. 注意，z 的 Euclid 模 $\|z\|$ 是一个非线性函数，所以(13.156)是非线性约束.

从 Farkas 引理和定理 3.2 我们知道，$Z^1(x^k)\cap Z^2(x^k)=\emptyset$（即不存在从 x^k 出发的能行下降方向）的充要条件为存在向量 $\lambda\geqslant 0$ 和 $\mu\in R^m$，使得

$$\nabla f(x^k)=\sum_{i=1}^{m}\mu_i a^i+\sum_{i\in I(x^k)}\lambda_i a^i, \tag{13.157}$$

其中 $a^i=(a_{i1},\ \cdots,\ a_{in})^T$. 换言之，能够找到一个从 x^k 出发的能行下降方向的充要条件为 $\nabla f(x^k)$ 不在相应于积极约束的向量 a^i 所张成的那个锥之中. 设后面的条件成立，令 t^{k+1} 是锥中最接近 $\nabla f(x^k)$ 的点，即

$$(t^{k+1})^T[t^{k+1}-\nabla f(x^k)]=0. \tag{13.158}$$

我们现在证明，(SFD)之解即最速能行下降向量由下式给出：

$$z^{k+1}=\frac{t^{k+1}-\nabla f(x^k)}{\|t^{k+1}-\nabla f(x^k)\|}. \tag{13.159}$$

定理 13.3[30]

令 z^{k+1} 由(13.159)给出，则它是求最速能行下降方向的问题(SFD)的最优解.

【证明】 对某个 μ 和 $\lambda\geqslant 0$，可以写

$$t^{k+1}=\sum_{i=1}^{m}\mu_i a^i+\sum_{i\in I(x^k)}\lambda_i a^i. \tag{13.160}$$

令 z 是满足(13.156)的任意能行方向，则

$$(t^{k+1})^T z=\sum_{i=1}^{m}\mu_i (a^i)^T z+\sum_{i\in I(x^k)}\lambda_i (a^i)^T z\geqslant 0. \tag{13.161}$$

从(13.158)和(13.159)推得

$$\begin{aligned}(z^{k+1})^T\nabla f(x^k)&=\frac{(t^{k+1}-\nabla f(x^k))^T\nabla f(x^k)}{\|t^{k+1}-\nabla f(x^k)\|}\\&=-\|t^{k+1}-\nabla f(x^k)\|,\end{aligned} \tag{13.162}$$

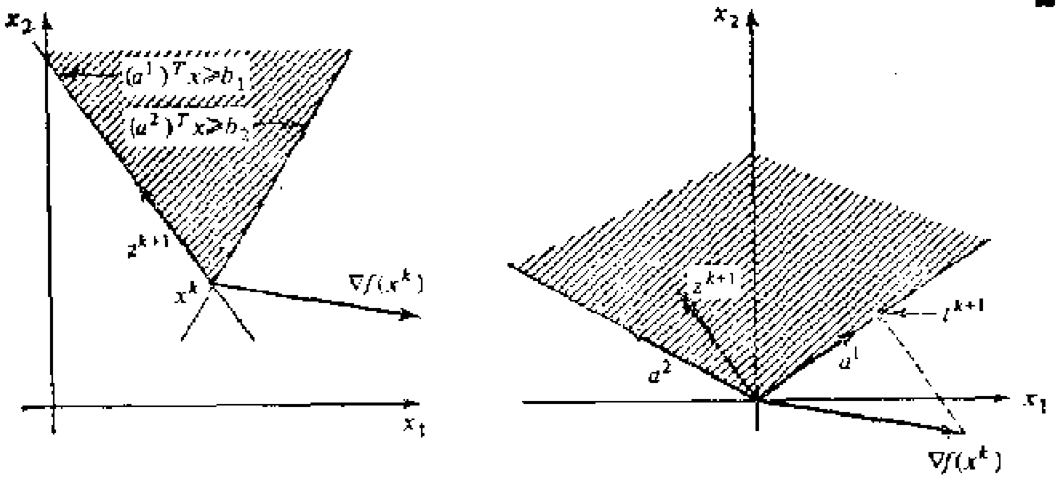
所以由(13.159)和(13.161)，有

$$z^T\nabla f(x^k)\geqslant -z^T(\|t^{k+1}-\nabla f(x^k)\|z^{k+1}). \tag{13.163}$$

利用 Schwarz 不等式[24]和(13.162)推得

$$z^T\nabla f(x^k) \geqslant -\|z\|\cdot\|t^{k+1}-\nabla f(x^k)\|\cdot\|z^{k+1}\| \tag{13.164}$$
$$\geqslant -\|t^{k+1}-\nabla f(x^k)\| \tag{13.165}$$
$$\geqslant (z^{k+1})^T\nabla f(x^k). \tag{13.166}$$

∎



(a) (b)

图 **13.3** 最速能行方向

最速能行方向的思想在图 13.3 中说明. (a) 中图表明由两个线性不等式约束定义的能行集，也表明两个约束都是积极的点 x^k 和该点的梯度. (b) 表明由向量 a^1 和 a^2 所张成的锥 $Z^1(x^k)$，也说明了 z^{k+1} 的构造.

z^{k+1} 的几何解释和定理 13.3 有一直接应用. 求最速能行下降方向的问题，可用一类非常特殊的二次规划来解，在这规划中要求的是由积极约束张成的锥到点 $\nabla f(x^k)$ 的最小距离. 这规划为

$$\min_{\lambda,\mu}\left\|\nabla f(x^k)-\sum_{i=1}^{m}\mu_i a^i-\sum_{i\in I(x^k)}\lambda_i a^i\right\|^2 \tag{13.167}$$

受限制于

$$\lambda_i\geqslant 0,\quad i\in I(x^k). \tag{13.168}$$

这里我们注意，这二次规划的求解不完全是平凡的，但有合用的有效解法，例如见 [2, 9, 11].

在选择一个能行下降方向时，求解非线性(二次)规划的要求可以避免，只要用限制方向向量各分量的别种约束代替 Euclid 模约束(13.156). 例如，我们可以用

$$\max_j |z_j|\leqslant 1 \tag{13.169}$$

来限制向量 z 的 L_∞(Chebyshev) 模，或者等价地，

$$-1 \leqslant z_j \leqslant 1, \quad j=1, \cdots, n. \tag{13.170}$$

用(13.170)代替(13.156)，产生了另一个求最速能行方向的问题，它是一个有特殊结构的线性规划. 这类附加的限界约束可以在[44]中找到.

有可能把上面概述的选取方向的步骤扩充为类似于有约束变尺度算法的方法，只要在公式中用左乘 H_k 的 $\nabla f(x^k)$ 代替 $\nabla f(x^k)$，这里 H_k 是 f 的 Hesse 逆阵的近似. 另外可以用 $(z^T B z)^{1/2}$ 这样的非 Euclid 模代替 Euclid 模，这里 B 是正定的，从一个迭代到另一个迭代时它可以变化.

在选定一个能行下降方向后，我们必须选步长 α_{k+1}^*，所以定义

$$\alpha_{k+1}^* = \min\{\alpha_{k+1}', \ \alpha_{k+1}''\}, \tag{13.171}$$

其中

$$\alpha_{k+1}' = \max\{\alpha_{k+1}: (z^{k+1})^T \nabla f(x^k + \alpha_{k+1} z^{k+1}) \leqslant 0\}, \tag{13.172}$$

$$\alpha_{k+1}'' = \max\{\alpha_{k+1}: (x^k + \alpha_{k+1} z^{k+1}) \in X\}, \tag{13.173}$$

对于 $\alpha_{k+1}' < +\infty$，我们有

$$\nabla f(x^k + \alpha_{k+1}' z^{k+1}) = 0.$$

令

$$x^{k+1} = x^k + \alpha_{k+1}^* z^{k+1},$$

我们回到选择方向的步骤. 看来似乎是，当给定一个初始能行点后，按这方式继续做，就有了一个完整的算法. 不幸，这样的方法不仅有非常慢的线性收敛速率，正象无约束最优化中的最速下降法一样，而且由于所谓锯齿现象[41]或拥塞现象[40]，约束可能阻止向最优点收敛. 这种现象可以用 Wolfe[39] 的例子来说明.

例 13.4.1

设我们有

$$\min f(x) = \frac{4}{3}[(x_1)^2 - x_1 x_2 + (x_2)^2]^{3/4} - x_3, \tag{13.174}$$

受限制于 $x \geqslant 0$. 函数 f 是凸的，所以是一个凸规划问题. 我们用前面的算法解这问题，从 $x^0 = \left(0, \dfrac{1}{4}, \dfrac{1}{2}\right)$ 出发. 在这点，约束

$x_1\geqslant 0$ 是积极的，且

$$-\nabla f(x^0)=\left(\left(\frac{1}{4}\right)^{1/2},\ -2\left(\frac{1}{4}\right)^{1/2},\ 1\right)^T. \tag{13.175}$$

因为 $-\nabla f(x^0)$ 是能行下降方向，我们沿着它移动，到达

$$x^1=\left(\frac{1}{2}\cdot\frac{1}{4},\ 0,\ \frac{1}{2}+\frac{1}{2}\left(\frac{1}{4}\right)^{1/2}\right), \tag{13.176}$$

在这里约束 $x_2\geqslant 0$ 变成积极的．按这方式继续，我们得到

$$x^2=\left(0,\ \frac{1}{4}\cdot\frac{1}{4},\ \frac{1}{2}+\frac{1}{2}\left(\frac{1}{4}\right)^{1/2}+\frac{1}{2}\left(\frac{1}{2}\cdot\frac{1}{4}\right)^{1/2}\right), \tag{13.177}$$

$x_1\geqslant 0$ 又变成积极的．由这方法产生的点在两个积极约束 $x_1\geqslant 0$ 和 $x_2\geqslant 0$ 之间摇摇摆摆，最后收敛（拥塞地）到点 $\hat{x}=(0,\ 0,\ \hat{x}_3)$，其中

$$\hat{x}_3=\frac{1}{2}+\left(1+\frac{\sqrt{2}}{2}\right)\left(\frac{1}{4}\right)^{1/2}. \tag{13.178}$$

要求读者说明 $\hat{x}$ 不是最优点． ▌

能行方向算法，包括"防止锯齿"过程，是由 McCormick[19]，Polak[26]，Zangwill[40]，Zoutendijk[41] 和其他人提出的．我们用 Zoutendijk 的一个方法结束本节．在这方法中，一旦某约束变成积极时，除非强迫它离开，否则它就不再离开积极约束集，这样就可以防止锯齿现象．

1．设 $x^0\in X$，计算 $\nabla f(x^0)$．令 $I_0=K_0=\emptyset$；令 $k=0$，并选一个小的数 $\varepsilon>0$．

2．用任何选取能行下降方向的方法求 z^{k+1}，它具有有界模长，且满足

(i) $$(a^i)^T z^{k+1}=0,\quad i=1,\ \cdots,\ m, \tag{13.179}$$

$$(a^i)^T z^{k+1}\begin{cases}=0, & i\in I_k,\\ \geqslant 0, & i\in I(x^k),\ i\overline{\in} I_k.\end{cases} \tag{13.180}$$

(ii) $$(\nabla f(x^{k+1})-\nabla f(x^k))^T z^{k+1}=0,\quad h\in K_k. \tag{13.181}$$

(iii) $$(z^{k+1})^T\nabla f(x^k)\leqslant -\varepsilon. \tag{13.182}$$

3．如果求方向的问题是不相容的，即没有 z^{k+1} 满足 (i) 至

(iii) 时，则

(i) 若 $I_k \cup K_k \neq \emptyset$，就从 I_k 或 K_k 中丢掉最旧的指标，进到步骤 2.

(ii) 若 $I_k \cup K_k = \emptyset$，则停止；最优解已得到.

4. 按照(13.171)给出的公式决定 α_{k+1}^*. 若 $\alpha_{k+1}^* = +\infty$，则停止，解是无界的.

5. 计算

$$x^{k+1} = x^k + \alpha_{k+1}^* z^{k+1} \tag{13.183}$$

和 $\nabla f(x^{k+1})$.

6. (i) 若 $\alpha_{k+1}^* = \alpha_{k+1}'$，则令 $I_{k+1} = I_k$ 和 $K_{k+1} = K_k \cup \{k\}$.

(ii) 若 $\alpha_{k+1}^* = \alpha_{k+1}''$，则令

$$I_{k+1} = I_k \cup \{i: (a^i)^T x^{k+1} = 0 \text{ 且 } (a^i)^T x^k > 0,\ i = m+1,\ \cdots,\ p\}, \tag{13.184}$$

并令 $K_{k+1} = \emptyset$.

(iii) 若 $\alpha_{k+1}' = \alpha_{k+1}''$，则按步骤 6(ii) 进行，但令 $K_{k+1} = \{k\}$.

7. 令 $k = k+1$，进到步骤 2.

注意，(13.181) 已经被加进来以获得共轭方向(见第 10 章)，直到出现下述情况：或者是最优解已求得，即对所有能行的 z，有

$$(z)^T \nabla f(x^k) > -\varepsilon; \tag{13.185}$$

或者积极约束集发生改变. 实际上，ε 是渐渐地缩减到一个非常小的正数，以保证几乎有最优性. 对于二次目标函数，每次迭代取 $\varepsilon = 0$.

如果没有初始能行点可用，我们可用线性规划的"第一相"过程解一个辅助线性规划：

$$\min \sum_{i=1}^{p} \xi_i \tag{13.186}$$

受限制于

$$\sum_{j=1}^{n} a_{ij} x_j + \xi_i = b_i, \quad i = 1,\ \cdots,\ m, \tag{13.187}$$

$$\sum_{j=1}^{n} a_{ij} x_j + \xi_i \geq b_i, \quad i = m+1,\ \cdots,\ p, \tag{13.188}$$

$$\xi \geqslant 0. \qquad (13.189)$$

如果这线性规划的最优解是 (x^*, ξ^*)，$\xi^*=0$，那末令 $x^0=x^*$，关于非线性目标函数的能行方向算法就可以开始.

Best、Ritter[4] 曾利用条件 (13.181) 达到接近共轭方向. 他们选择方向的步骤是基于 Frank-Wolfe[12] 方法的类似部分，在那里，要极小化的是目标函数的一个线性近似，而受限制于能确保搜索方向能行性的约束.

最后，提一下另两个能行方向法，它们是无约束极小化方法的扩充，用来求解具有线性约束的非线性规划问题. 第一个方法是 Ritter[31] 导出的，它对一大类问题是收敛的，且对某些严格凸函数有超线性的收敛速率. 这方法是同一作者仅用一阶导数的无约束问题算法的扩充，在第 10 章曾提到过. 第二个方法属于 McCormick[20]，既用一阶导数也用二阶导数，收敛到满足最优性二阶必要条件的点.

13.5 非线性约束的投影法和能行方向法

在第 13.3 节和 13.4 节中的梯度投影算法和能行方向算法，其大部分已被扩充到具有非线性约束的数学规划上. 但是，用这些方法求解这种问题，比线性约束的问题要困难得多. 如果约束是线性的，对目标函数进行一维搜索的方向或者在能行集的内部，或者在相应于积极约束集的诸超平面的交集中. 在两种情形下，从当前的点沿着搜索方向移动一个小距离将不会有脱离能行集的危险. 扩充投影法或能行方向法到非线性约束情形，由此引起的困难是，搜索方向限制在积极的非线性约束的切平面之交集中，这时，沿这搜索方向的任何移动可能导致非能行点，就需要额外的步骤使之回到能行集. 重复地沿搜索方向移动并回到能行集，产生了图 13.4 所示的“花边”现象.

积极约束的这种直接处理使这算法可观地慢下来. 由于这原因，利用惩罚函数法间接处理非线性约束似乎更有利，除非非线性约束有某种特殊类型或结构，如几何规划中的那样. 下一章给出

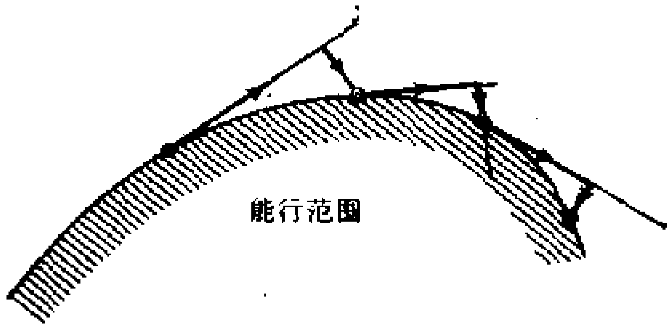


图 13.4 沿一个积极非线性约束形成的"花边"现象

的方法也可以推荐来解某些非线性约束问题，虽然这些算法的收敛速率一般是慢的. 我们在结束这一章之前，还需对无约束最优化方法在非线性约束数学规划问题中的某些扩充作出简要的评论.

考虑一般非线性规划

$$\min f(x) \tag{13.190}$$

受限制于

$$g_i(x) \geqslant 0, \quad i=1, \cdots, m, \tag{13.191}$$

$$h_j(x)=0, \quad j=1, \cdots, p, \tag{13.192}$$

其中 f, g_i, h_j 是连续可微函数，并且至少有一个约束是非线性的. 求解这个问题的一个方法是 Goldfarb 算法的扩充，它是属于 Davies[6] 的.

在点 x^k 线性化了的不等式约束 $g_i(x)\geqslant 0$ 由下式给出:

$$g_i(x^k)+(x-x^k)^T \nabla g_i(x^k) \geqslant 0, \tag{13.193}$$

类似地，线性化了的等式约束 $h_j(x)=0$ 是

$$h_j(x^k)+(x-x^k)^T \nabla h_j(x^k)=0. \tag{13.194}$$

如果 $g_i(x^k)=0$, 则线性化约束(13.193)定义了一个闭半空间，它是由曲面 $g_i(x)=0$ 在 x^k 处的切超平面产生的. 若 $h_j(x^k)=0$, 则满足(13.194)的点位于 $h_j(x)=0$ 在 x^k 处的切超平面中.

整个计算过程中，所有等式约束必须保持"积极". 一个非线性约束称为在 x^k 是积极的，如果相应的线性约束在 x^k 处是积极的. 如果任意 g_i 或 h_j 都是线性的，那就按 Goldfarb 算法的同样方式来处理. 给定一个能行点 x^k, 决定 x^k 处的积极约束集并计算

搜索方向 z^{k+1}，它是适当地投影到相应于积极约束的超平面的交集中．下一步是决定沿着 z^{k+1} 方向步长的一个界限 $\bar{\alpha}_{k+1}$．求出沿 z^{k+1} 到最近的线性化约束的距离，记这距离为 θ，则令

$$\bar{\alpha}_{k+1}=\min\{c,\ \theta\} \tag{13.195}$$

其中 c 是某个正的常数，例如 $c=2$．这样的 $\bar{\alpha}_{k+1}$ 将限制步长并避免从能行点移动太远．现在我们令

$$x^{k,1}=x^{k}+\bar{\alpha}_{k+1}z^{k+1}, \tag{13.196}$$

并求在 $x^{k,1}$ 处的所有约束的值．如果任何约束被违反，就决定一个指向能行集的返回方向向量 r^k．沿着这方向作一移动，得到一内点 $x^{k,2}$，按一个内插格式在能行集的边界上确定一点 $x^{k,3}$．为了这目的，一个有趣的内插方法是由 Beamer、Wilde[3] 建立的．然后沿着连结 x^k 和 $x^{k,3}$ 的线段求 f 的极小值点．如果这样的极小值点 $x^{k,4}$ 已求得，并且如果能行集是凸的，那末 $x^{k,4}$ 是在能行集的内部，我们必须沿着例如最速下降方向再次移动到边界，达到新的一点 x^{k+1}．如果在 x^k 和 $x^{k,3}$ 之间没有找到极小值点，新的搜索就从 $x^{k,3}$ 出发沿 z^{k+1} 方向开始．图 13.5 表明了一种典型的情况．

为了在一个可用的算法中实现 Davies 方法，很多技术性的细

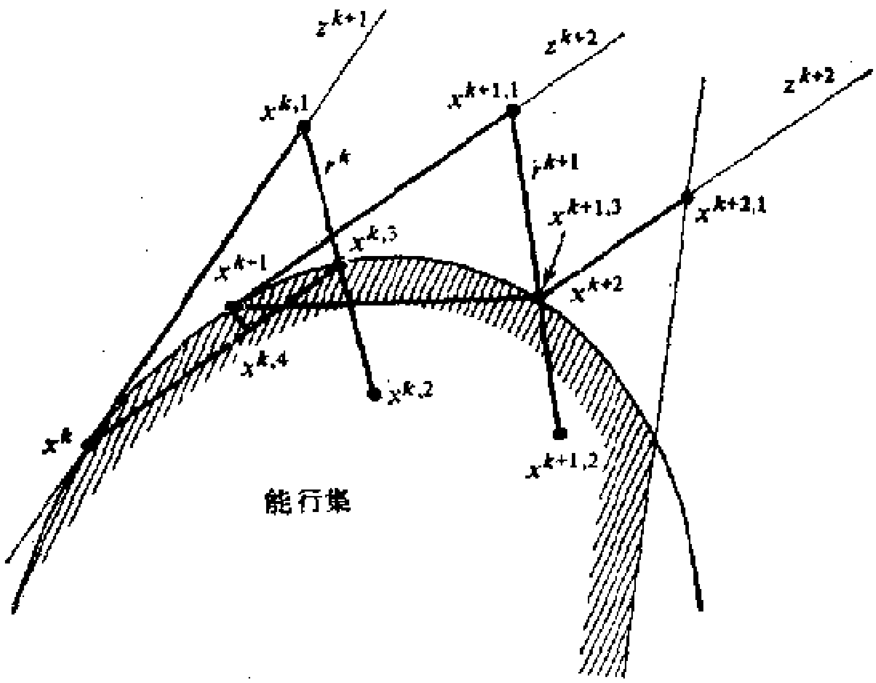


图 **13.5**　用 Davies 方法沿积极的非线性约束移动的情况

节是重要的. 关于细节，读者可参看 Davies[6] 和 Rosen[34].

现在转到 Zoutendijk[41,42] 为非线性约束导出的能行方向法. 考虑问题

$$\min f(x) \tag{13.197}$$

受限制于

$$g_i(x) \geqslant 0, \quad i=1, \cdots, m, \tag{13.198}$$

其中 f 和 g_i 是连续可微的. 对于任何能行点 x 和 $\varepsilon \geqslant 0$, 定义

$$X_\varepsilon(x) = \{i: g_i(x) \leqslant \varepsilon\}, \tag{13.199}$$

并令 X 是能行集. 在第 k 次迭代, 给定了能行点 x^k 和 $\varepsilon^k > 0$, 我们求解如下的“选取方向的线性规划”

$$\min \sigma \tag{13.200}$$

受限制于

$$\sigma \geqslant z^T \nabla f(x^k), \tag{13.201}$$

$$z^T \nabla g_i(x^k) \geqslant \sigma, \quad i \in X_{\varepsilon^k}(x^k), \tag{13.202}$$

$$|z_j| \leqslant 1, \quad j=1, \cdots, n. \tag{13.203}$$

令最优解是 $\sigma^{k+1} \in R$, $z^{k+1} \in R^n$. 如果 $\sigma^{k+1} \leqslant -\varepsilon^k$, 令 $\varepsilon^{k+1} = \varepsilon^k$; 否则令 $\varepsilon^{k+1} = \dfrac{1}{2}\varepsilon^k$. 如(13.171)至(13.173)给出的那样计算 α^*_{k+1}, 并令 $x^{k+1} = x^k + \alpha^*_{k+1} z^{k+1}$. Zangwill[40] 证明这方法收敛到(13.197)和(13.198)给出的问题的解，在[40]中他称 x^* 为解，如果从 x^* 出发找不到能行下降方向的话. 当然，α^*_{k+1} 的计算可能是非常繁复的，要求设计几个迭代，象 Davies 方法中概述的那些.

惩罚函数法和线性约束非线性规划的算法的成功实现，推动了两个新近的工作，在这两个工作中，非线性约束规划是用一个线性约束规划的序列来求解的，其中利用了惩罚函数-Lagrange 式方法.

再次考虑不等式约束问题

$$\min f(x) \tag{13.204}$$

$$g_i(x) \geqslant 0, \quad i=1, \cdots, m, \tag{13.205}$$

其中 f 和 g_i 分别是 $x \in R^n$ 的连续可微的凸函数和凹函数. Rosen、

Kreuser[35,36] 的方法可描述如下：从任意点 $x^0\in R^n$ 开始．在第 k 次迭代，我们有一个指标集 $I(x^k)$，定义为

$$I(x^k)=\{i:g_i(x^k)\leqslant 0,\ i=1,\ \cdots,\ m\}. \tag{13.206}$$

按下式计算 λ^k：

$$\lambda^k=J^+(x^k)\nabla f(x^k), \tag{13.207}$$

其中

$$J^+(x^k)=[(J(x^k))^TJ(x^k)]^{-1}(J(x^k))^T, \tag{13.208}$$

而 $J(x^k)$ 是 $i\in I(x^k)$ 的 $g_i(x^k)$ 在 x^k 的 Jacobi 矩阵，即其列是 $\nabla g_i(x^k)$ 的矩阵，$i\in I(x^k)$．$J^+(x^k)$ 是 $J(x^k)$ 的广义逆．矩阵 $(J(x^k))^TJ(x^k)$假定是非异的．其次求解线性约束问题

$$\min\{f(x)-\sum_{\lambda_i^k>0}\lambda_i^k[g_i(x)-g_i(x^k)-(x-x^k)^T\nabla g_i(x^k)]\} \tag{13.209}$$

受限制于

$$g_i(x^k)+(x-x^k)^T\nabla g_i(x^k)\geqslant 0,\quad i=1,\cdots,\ m, \tag{13.210}$$

在求解时，不妨用 Goldfarb 方法．设解为 x^{k+1}．如果 x^{k+1} 在某容许误差之内满足问题(13.204)和(13.205)的一阶 Kuhn-Tucker 条件，就停止；否则令 $k=k+1$，重复前面的步骤．可以证明，只要满足某些正则性假定，这个方法具有二次收敛速率．注意，线性约束决不会出现在(13.209)中，但包含在(13.210)中．

Robinson[32] 提出用一个类似的方法，来求解具有下面附加等式约束的问题(13.204)和(13.205)：

$$h_j(x)=0,\quad j=1,\ \cdots,\ p, \tag{13.211}$$

其中 h_j 和 f, g_i 是二次连续可微的．解法如下：在第 k 次迭代，我们有向量 $x^k\in R^n$ 和 $\lambda^k\in R^m$, $\lambda^k\geqslant 0$, $\mu^k\in R^p$，解线性约束问题

$$\min_x\left\{f(x)-\sum_{i=1}^{m}\lambda_i^k[g_i(x)-\tilde{g}_i(x^k,\ x)]\right.$$
$$\left.-\sum_{j=1}^{p}\mu_j^k[h_j(x)-\tilde{h}_j(x^k,\ x)]\right\} \tag{13.212}$$

受限制于

$$\tilde{g}_i(x^k,\ x)=g_i(x^k)+(x-x^k)^T\nabla g_i(x^k)\geqslant 0,\ \ i=1,\ \cdots,\ m, \tag{13.213}$$

$$\tilde{h}_j(x^k, x) = h_j(x^k) + (x - x^k)^T \nabla h_j(x^k) = 0, \quad j = 1, \cdots, p. \tag{13.214}$$

令解是 x^{k+1}. 得到这线性约束问题的相应 Lagrange 乘子 $\lambda^{k+1} \geqslant 0$, μ^{k+1}. 如果有多于一个这样的三重组$(x^{k+1}, \lambda^{k+1}, \mu^{k+1})$，从中选取在某种模下最接近$(x^k, \lambda^k, \mu^k)$的一个. 如果$(x^{k+1}, \lambda^{k+1}, \mu^{k+1})$在某容许误差内满足规划 (13.204), (13.205) 和 (13.211) 的一阶 Kuhn-Tucker 条件，就停止；否则令 $k=k+1$, 重新求解(13.212)到(13.214). 开始这算法的点 x^0 不必是能行点. 再则，在关于定义这规划的函数的适当假定下，这算法二次收敛到最优点.

练 习

13.A. 设由(13.1)和(13.2)定义的 $x \in R^n$ 的集是凸的，说明复合形法的移动是很好地定义的. 给一个 R^2 中问题的例子，其中某些移动可以失败[提示：试试弯曲的山谷]. 说明无约束单纯形法可用来求复合形法的初始能行点 x^0. 用单纯形法求这样一点如果失败，是否证明能行集是空集?

13.B. 对例 9.3.1 并具有约束

$$-2x_1 + x_2 + 2 \geqslant 0 \tag{13.215}$$

的问题，应用 Rosenbrock 旋转方向法的有约束形式，进行两个勘探阶段，比较用这方法得到的最好点 x_B^2 与真正的约束极小值点. 用例 9.3.1 中同样参数并令 $\varepsilon = 10^{-4}$.

13.C. 继续例 13.2.2 如下：首先丢掉约束(13.45)，并按(13.50)从 P_2 计算新的投影矩阵 P_1. 然后增加约束

$$\frac{3}{5}x_2 - \frac{4}{5}x_3 = 2, \tag{13.216}$$

并计算新矩阵 P_2. 用 (13.27), (13.28) 和递推关系 (13.35) 至 (13.43).

13.D. 用一般线性约束非线性规划的 Goldfarb 方法求解下述问题：

$$\min f(x) = \frac{3}{2}(x_1)^2 + \frac{1}{2}(x_2)^2 - x_1x_2 - 2x_1 \tag{13.217}$$

受限制于

$$x_1 + x_2 = 5, \tag{13.218}$$

$$6x_1 - x_2 \geqslant -12, \tag{13.219}$$

$$-x_1+x_2\geqslant 3, \tag{13.220}$$

从点 $x^0=(-1, 6)$ 开始计算，注意约束系数尚未规范化.

13.E. 证明关系式(13.92)，(13.94)和(13.95).

13.F. 参考 Murtagh、Sargent 的著作[22]，画出他们的梯度投影法的框图. 对于下面的问题，比较由他们的算法产生的点与 Goldfarb 算法产生的点：

$$\min (x_1)^2+4(x_2)^2 \tag{13.221}$$

受限制于

$$x_1+2x_2\geqslant 1, \tag{13.222}$$

$$x_1-x_2\geqslant 0, \tag{13.223}$$

$$x_1\geqslant 0, \tag{13.224}$$

从点 $x^0=(8, 8)$ 开始.

13.G. 用 13.3 节的算法重新求解练习 13.D 和 13.F 中给出的二次规划. 比较现在所经过的点与这些练习中得到的相应点.

13.H. 证明：$z^{k+1}\in E^n$ 是从点 x^k 出发的、满足 (3.110) 和 (3.111) 的能行下降方向的充要条件为，它是由 (13.152) 和 (13.153) 给出的 $Z^1(x^k)\cap Z^2(x^k)$ 之元素. 给出线性约束非线性规划的一个例子，其中从点 x^k 出发没有能行下降方向，但这并不蕴涵 x^k 的最优性. 找一类问题，对于它们，这种最优性是有保证的.

13.I. 注意，找最速能行下降方向的问题 (SFD) 是一个凸规划. 用第 5 章的结果导出它的对偶规划问题. 与(13.167)和(13.168)比较.

13.J. 证明例 13.4.1 中的目标函数是凸的. 并证明 $\hat{x}$ 不是最优的. 在这问题中最优解是什么？

13.K. Levitin、Polyak[18] 提出了凸规划的梯度投影能行方向算法，在这方法中第 k 次迭代时，能行点 x^k 已给定，通过求解二次规划

$$\min_{y\in X}(y-(x^k-\nabla f(x^k)))^T(y-(x^k-\nabla f(x^k))), \tag{13.225}$$

将方向 $x^k-\nabla f(x^k)$ 投影在能行集 X 上. 如果 y^k 是上述规划的最优解，则搜索方向 z^{k+1} 为

$$z^{k+1}=y^k-x^k. \tag{13.226}$$

(a) 如果

$$X=\{x: Ax=b\}, \tag{13.227}$$

其中 A 是秩为 m 的 $m\times n$ 阵. 证明，y^k 可以明显地写成 A, x^k 和 $\nabla f(x^k)$ 的函数. [提示：用(13.94).]

(b) 如果

$$X=\{x: a_j \leqslant x_j \leqslant b_j,\ j=1, \cdots, n\}, \tag{13.228}$$

其中 a_j 和 b_j 是给定的界限. 求 y^k 的显表达式.

提示: 考虑三种不同情形

(i)
$$x_j^k-\frac{\partial f(x^k)}{\partial x_j} \leqslant a_j, \tag{13.229}$$

(ii)
$$a_j<x_j^k-\frac{\partial f(x^k)}{\partial x_j}<b_j, \tag{13.230}$$

(iii)
$$b_j \leqslant x_j^k-\frac{\partial f(x^k)}{\partial x_j}. \tag{13.231}$$

13.L. 给定非线性规划:

$$\min (x_1-4)^2+(x_2-4)^2 \tag{13.232}$$

受限制于

$$3(x_1)^2+(x_2)^2-2x_1x_2-4x_1 \leqslant 12, \tag{13.233}$$

$$3x_1+4x_2 \leqslant 28. \tag{13.234}$$

(a) 用 Davies 方法解这问题, 从点 $x^0=(2, 0)$ 开始, 将这个算法进行几步, 建立一个决定返回方向 r^k 的规则. 参考 Davies 的著作[6], 与他的建议比较.

(b) 从同一点 x^0 开始, 对同样的问题, 试试 Zoutendijk 方法.

(c) 用 Rosen-Kreuser 和 Robinson 方法去解这个问题, 进行几步. 画出用每种方法获得的点的行程. 对于线性约束的子问题, 用 Goldfarb 算法.

参 考 文 献

1. BARTELS, R. H., G. H. GOLUB, and M. A. SAUNDERS, "Numerical Techniques in Mathematical Programming," in *Nonlinear Programming*, J. B. Rosen, O. L. Mangasarian, and K. Ritter (Eds.), Academic Press, New York, 1970.
2. BEALE, E. M. L., "Numerical Methods," in *Nonlinear Programming*, J. Abadie (Ed.), North-Holland Publishing Co., Amsterdam, 1967.
3. BEAMER, J. H., and D. J. WILDE, "A Minimax Search Plan for Constrained Optimization Problems," *J. Optimization Theory & Appl.*, **12**, 439–446 (1973).
4. BEST, M. J., and K. RITTER, "An Accelerated Conjugate Direction Method to Solve Linearly Constrained Minimization Problems," Research Report CORR 73-16, Department of Combinatorics and Optimization, University of Waterloo, Ontario, Canada, August 1973.
5. BOX, M. J., "A New Method of Constrained Optimization and a Comparison with Other Methods," *Computer J.*, **8**, 42–52 (1965).

6. Davies, D., "Some Practical Methods of Optimization," in *Integer and Nonlinear Programming*, J. Abadie (Ed.), North-Holland Publishing Co., Amsterdam, 1970.

7. Davies, D., and Swann, W. H., "Review of Constrained Optimization," in *Optimization*, R. Fletcher (Ed.), Academic Press, London, 1969.

8. Fletcher, R., "The Calculation of Feasible Points for Linearly Constrained Optimization Problems," Report R6354, A.E.R.E. Harwell, United Kingdom, 1970.

9. Fletcher, R., "A General Quadratic Programming Algorithm," *J. Inst. Math. Appl.*, **7**, 76–91 (1971).

10. Fletcher, R., "Minimizing General Functions Subject to Linear Constraints," in *Numerical Methods for Non-linear Optimization*, F. A. Lootsma (Ed.), Academic Press, London, 1972.

11. Fletcher, R., and M. P. Jackson, "Minimization of a Quadratic Function of Many Variables Subject Only to Lower and Upper Bounds," Report T.P. 528, A.E.R.E. Harwell, United Kingdom, May 1973.

12. Frank, M., and P. Wolff, "An Algorithm for Quadratic Programming," *Naval Res. Log. Quart.*, **3**, 95–110 (1956).

13. Glass, H., and L. Cooper, "Sequential Search: A Method for Solving Constrained Optimization Problems," *J. Assoc. Comp. Mach.*, **12**, 71–82 (1965).

14. Goldfarb, D., "Extension of Davidon's Variable Metric Method to Maximization under Linear Inequality and Equality Constraints," *SIAM J. Appl. Math.*, **17**, 739–764 (1969).

15. Goldfarb, D., "Extensions of Newton's Method and Simplex Methods for Solving Quadratic Programs," in *Numerical Methods for Non-linear Optimization*, F. A. Lootsma (Ed.), Academic Press, London, 1972.

16. Goldfarb, D., and L. Lapidus, "Conjugate Gradient Method for Nonlinear Programming Problems with Linear Constraints," *Ind. & Eng. Chem. Fund.*, **7**, 142–151 (1968).

17. Householder, A. S., "Unitary Triangularization of a Nonsymmetric Matrix," *J. Assoc. Comp. Mach.*, **5**, 339–342 (1958).

18. Levitin, E. S., and B. T. Polyak, "Constrained Minimization Methods," *USSR Comp. Maths. & Math. Phys.*, **6**, 5, 1–50 (1966).

19. McCormick, G. P., "Anti-zigzagging by Bending," *Management Science*, **15**, 315–320 (1969).

20. McCormick, G. P., "A Second Order Method for the Linearly Constrained Nonlinear Programming Problem," in *Nonlinear Programming*, J. B. Rosen, O. L. Mangasarian, and K. Ritter (Eds.), Academic Press, New York, 1970.

21. Mitchell, R. A., and J. L. Kaplan, "Nonlinear Constrained Optimization by a Nonrandom Complex Method," *J. Research of the NBS-C. Eng. and Instr.*, **72C**, 249–258 (1968).

22. MURTAGH, B. A., and R. W. H. SARGENT, "A Constrained Minimization Method with Quadratic Convergence," in *Optimization*, R. Fletcher (Ed.), Academic Press, London, 1969.

23. NELDER, J. A., and R. MEAD, "A Simplex Method for Function Minimization," *Computer J.*, **7**, 308–313 (1965).

24. NOBLE, B., *Applied Linear Algebra*, Prentice-Hall, Englewood Cliffs, N.J., 1969.

25. POLAK, E., "On the Convergence of Optimization Algorithms," *Rev. Inform. Rech. Oper.*, **16**, 17–34 (1969).

26. POLAK, E., *Computational Methods in Optimization*, Academic Press, New York, 1971.

27. POWELL, M. J. D., "Quadratic Termination Properties of Minimization Algorithms I. Statement and Discussion of Results," *J. Inst. Math. Appl.*, **10**, 333–342 (1972).

28. POWELL, M. J. D., "Quadratic Termination Properties of Minimization Algorithms II. Proofs of Theorems," *J. Inst. Math. Appl.*, **10**, 343–357 (1972).

29. POWELL, M. J. D., "Unconstrained Minimization and Extensions for Constraints," Report T.P. 495, A.E.R.E. Harwell, United Kingdom, August 1972.

30. POWELL, M. J. D., "Introduction to Constrained Optimization," Report C.S.S.1, A.E.R.E. Harwell, United Kingdom, January 1974.

31. RITTER, K., "A Superlinearly Convergent Method for Minimization Problems with Linear Inequality Constraints," *Math. Prog.*, **4**, 44–71 (1973).

32. ROBINSON, S. M., "A Quadratically Convergent Algorithm for General Nonlinear Programming Problems," *Math. Prog.*, **3**, 145–156 (1972).

33. ROSEN, J. B., "The Gradient Projection Method for Nonlinear Programming, Part I. Linear Constraints," *J. SIAM*, **8**, 181–217 (1960).

34. ROSEN, J. B., "The Gradient Projection Method for Nonlinear Programming, Part II. Nonlinear Constraints," *J. SIAM*, **9**, 514–532 (1961).

35. ROSEN, J. B., and J. L. KREUSER, "A Gradient Projection Algorithm for Nonlinear Constraints," in *Numerical Methods for Non-linear Optimization*, F. A. Lootsma (Ed.), Academic Press, London, 1972.

36. ROSEN, J. B., and J. L. KREUSER, "A Quadratically Convergent Lagrangian Algorithm for Nonlinear Constraints," Technical Report No. 166, Computer Science Department, The University of Wisconsin, Madison, November 1972.

37. ROSENBROCK, H. H., "An Automatic Method for Finding the Greatest or Least Value of a Function," *Computer J.*, **3**, 175–184 (1960).

38. SPENDLEY, W., G. R. HEXT, and F. R. HIMSWORTH, "Sequential Application of Simplex Designs in Optimisation and Evolutionary Operation," *Technometrics*, **4**, 441–461 (1962).

39. WOLFE, P., "On the Convergence of Gradient Methods under Constraint," *IBM J. Res. Dev.*, **16**, 407–411 (1972).

40. ZANGWILL, W. I., *Nonlinear Programming: A Unified Approach*, Prentice-Hall, Englewood Cliffs, N.J., 1969.

41. ZOUTENDIJK, G., *Methods of Feasible Directions*, Elsevier Publishing Co., Amsterdam, 1960.

42. ZOUTENDIJK, G., "Nonlinear Programming: A Numerical Survey," *SIAM J. Control*, **4**, 194–210 (1966).

43. ZOUTENDIJK, G., "Nonlinear Programming, Computational Methods," in *Integer and Nonlinear Programming*, J. Abadie (Ed.), North-Holland Publishing Co., Amsterdam, 1970.

44. ZOUTENDIJK, G., "Some Algorithms Based on the Principle of Feasible Directions," in *Nonlinear Programming*, J. B. Rosen, O. L. Mangasarian, and K. Ritter (Eds.), Academic Press, New York, 1970.

第 14 章

近似型算法

第 12 章和 13 章介绍的算法显然是面向无约束最优化的，即它们或者把一个有约束问题变为无约束问题，或者通过适当调整搜索方向和步长，把无约束方法扩充到有约束问题．本章叙述的约束最优化方法按下述意义是面向线性规划的：其大部分步骤是求解线性规划的子问题，或者，当所考虑的问题恰是线性规划时，算法就成为著名的单纯形法的某种变形．线性规划是最优化理论的一个如此广泛而重要的分支，以致在大多数课程表中它通常是作为一门单独的课程来讲授，因而不能包含在本书中．然而，对于理解这里的论题，线性规划知识并非必需，读者只需知道单纯形法是解线性规划的一个极其有力的工具即可．关于线性规划及其某些扩充的一个基本研究，读者可以参看 Dantzig[12]．在那里还讨论了利用单纯形法的变形求解某些非线性规划问题．事实上，在解复杂而大型的线性规划时单纯形法的成就使得许多学者相信，大多数非线性规划能够而且应当通过本章的这一类方法来求解．我们将不叙述沿着这些线索所提出的所有方法．关于此处所给出方法的补充材料，例如可以在 Beale 的评论[7, 8]和那里提到的参考资料中找到．

第 14.1 节从一个算法开始，这算法基于对非线性规划的目标函数和约束函数的逐次线性化．然后提出一个方法，它适用于目标函数被二次函数逐次逼近的线性约束问题．既约梯度法类似于单纯形法，把变量分成基本的(相关的)和非基本的(独立的)，它将在 14.2 节中讨论．割平面算法是 14.3 节的论题，按下述意义它们不同于先前的算法：它们只适用于凸(或广义凸)问题，而最优解是沿着一条非能行的行程被接近的，这个行程通过求解一系列线性规划而得到．最后一节是关于一族非凸规划的解法，其中某些约束

描述凸集的非凸补集，还给出了一个详细的收敛性证明，这个证明也阐述了由 Zangwill[40]提出的处理收敛性结果的一个一般途径.

14.1 近似规划方法

用一系列线性规划去求解一般非线性规划问题，最简单的途径是，用线性(仿射)函数去近似问题中的非线性函数，然后去解所产生的线性规划. 一个直接的缺点是线性近似通常只是局部有效的，即在非线性函数被近似的那些点的周围有效，或者至多在作出近似的这些点的某个小邻域中有效. 所以，为了发展一个有用的算法，我们在当前点周围作函数的近似，且从这点所作的移动限于取小步长. 为了这个理由，本节描述的方法也称为"小步长"梯度法，以区别于第 13 章介绍的"大步长"方法.

这里首先给出的近似方法是 Griffith、Stewart[19] 导出的. 考虑带有有界变量的一般非线性规划

(BP)
$$\min f(x) \tag{14.1}$$

受限制于

$$g_i(x)\geqslant 0, \quad i=1,\ \cdots,\ m, \tag{14.2}$$

$$h_l(x)=0, \quad l=1,\ \cdots,\ p, \tag{14.3}$$

$$u_j\geqslant x_j\geqslant 0, \quad j=1,\ \cdots,\ n, \tag{14.4}$$

其中 f、g_i、h_l 是可微函数，且上界 u_j 是给定的. 如果下界不全为零，那么用一个简单的变量代换可以容易地将它们化成上述形式.

假设 $\bar{x}$ 是(BP)的一个能行解，在 $\bar{x}$ 周围将每个非线性函数线性化，得到

$$f(x)\cong\tilde{f}(x,\ \bar{x})=f(\bar{x})+(x-\bar{x})^T\nabla f(\bar{x}), \tag{14.5}$$

$$g_i(x)\cong\tilde{g}_i(x,\ \bar{x})=g_i(\bar{x})+(x-\bar{x})^T\nabla g_i(\bar{x}), \quad i=1,\ \cdots,\ m, \tag{14.6}$$

$$h_l(x)\cong\tilde{h}_l(x,\ \bar{x})=h_l(\bar{x})+(x-\bar{x})^T\nabla h_l(\bar{x}), \quad l=1,\ \cdots,\ p. \tag{14.7}$$

引入新的向量变量 y，它定义为

$$y=x-\bar{x}. \tag{14.8}$$

因为 y 的符号一般是没有限制的，我们利用公式

$$y_j=y_j^+-y_j^-,\quad j=1,\ \cdots,\ n,\tag{14.9}$$

其中

$$y_j^+\geqslant 0,\quad y_j^-\geqslant 0,\quad j=1,\ \cdots,\ n,\tag{14.10}$$

并且

$$y_j^+=\begin{cases}y_j, & y_j\geqslant 0,\\ 0, & y_j<0,\end{cases}\tag{14.11}$$

$$y_j^-=\begin{cases}-y_j, & y_j\leqslant 0,\\ 0, & y_j>0.\end{cases}\tag{14.12}$$

量 $|y_j|$ 代表从 $\bar{x}_j$ 到 x_j 的距离，$\bar{x}_j$ 是近似点 $\bar{x}$ 的分量．我们希望这个距离受限制于下列不等式：

$$|y_j|\leqslant m_j,\quad j=1,\ \cdots,\ n,\tag{14.13}$$

其中 m_j 是给定的常数．因为向量 x 也是有界的，结合(14.4)与(14.8)至(14.13)，得到

$$0\leqslant y_j^+\leqslant\min\{m_j,\ u_j-\bar{x}_j\},\tag{14.14}$$

$$0\leqslant y_j^-\leqslant\min\{m_j,\ \bar{x}_j\}.\tag{14.15}$$

近似于(BP)的线性规划为

$$\min\tilde{f}(x,\ \bar{x})=f(\bar{x})+(y^+)^T\nabla f(\bar{x})-(y^-)^T\nabla f(x)\tag{14.16}$$

受限制于

$$\tilde{g}_i(x,\ \bar{x})=g_i(\bar{x})+(y^+)^T\nabla g_i(\bar{x})-(y^-)^T\nabla g_i(\bar{x})\geqslant 0,\quad i=1,\ \cdots,\ m,\tag{14.17}$$

$$\tilde{h}_l(x,\ \bar{x})=h_l(\bar{x})+(y^+)^T\nabla h_l(\bar{x})-(y^-)^T\nabla h_l(\bar{x})=0,\quad l=1,\ \cdots,\ p,\tag{14.18}$$

还受限制于(14.14)和(14.15)．

Griffith 和 Stewart 的算法由一个迭代格式所组成，它从一点 $\bar{x}$ 开始，求解上述线性规划．如果这个近似规划的最优解 x_L^* 对于(BP)是能行的，那么把它作为下一个线性规划的新的近似点，这下一个规划具有与先前相同的界限．然而，如果 x_L^* 是非能行的，它不被接受，这时就用缩小了的 m_j 值重复求解原来的线性规划．

一个点是(BP)的最优点，如果它在某个允许范围内是能行的，并且关于 y 和 $f(x)$ 的相继值的某个收敛准则是满足的.

界限 m_j 的值对于算法的成功与否影响很大. 如果它们太小，步长也就取得小，进展会相当慢；另一方面，m_j 的值太大，可能得到非能行解，于是需要逐步缩小界限的值. 从一个近似点移到另一个点时，所有非线性函数要重新取近似，并且，通常必须去解一个全新的线性规划. 这个方法并不局限于凸规划或它的扩充. 然而，甚至对于凸规划，局部最优解的收敛性也没有得到证明，虽然在实践中这个方法在凸的情况下通常是收敛的. Griffith-Stewart 方法主要适用于这样的非线性化最优化问题：它有很多变量，但可能缺少凸性假定，然而除了线性约束外只有少量非线性约束. 在这种问题中常常要改善某些现有解，而理论上的收敛性考虑并不重要. 当然，要假设这个方法的使用者可以使用有效的线性规划程序.

例 14.1.1

为了说明这个算法，假设求解的非线性规划是

$$\min f(x) = -2x_1 - x_2 \tag{14.19}$$

受限制于

$$g_1(x) = 25 - (x_1)^2 - (x_2)^2 \geqslant 0, \tag{14.20}$$

$$g_2(x) = 7 - (x_1)^2 + (x_2)^2 \geqslant 0, \tag{14.21}$$

$$5 \geqslant x_1 \geqslant 0, \quad 10 \geqslant x_2 \geqslant 0. \tag{14.22}$$

我们从点 $\bar{x} = (2, 2)$ 开始，界限 m_j 取 $m_1 = 1.5$、$m_2 = 2.5$，非线性规划和第一次近似线性规划(关于原来的变量)示于图 14.1 中. 线性规划是

$$\min \tilde{f}(x, \bar{x}) = -6 - 2y_1^+ + 2y_1^- - y_2^+ + y_2^- \tag{14.23}$$

受限制于

$$\tilde{g}_1(x, \bar{x}) = 17 - 4y_1^+ + 4y_1^- - 4y_2^+ + 4y_2^- \geqslant 0, \tag{14.24}$$

$$\tilde{g}_2(x, \bar{x}) = 7 - 4y_1^+ + 4y_1^- + 4y_2^+ - 4y_2^- \geqslant 0, \tag{14.25}$$

$$1.5 \geqslant y_1^+ \geqslant 0, \tag{14.26}$$

$$1.5 \geqslant y_1^- \geqslant 0, \tag{14.27}$$

$$2.5 \geqslant y_2^+ \geqslant 0, \tag{14.28}$$

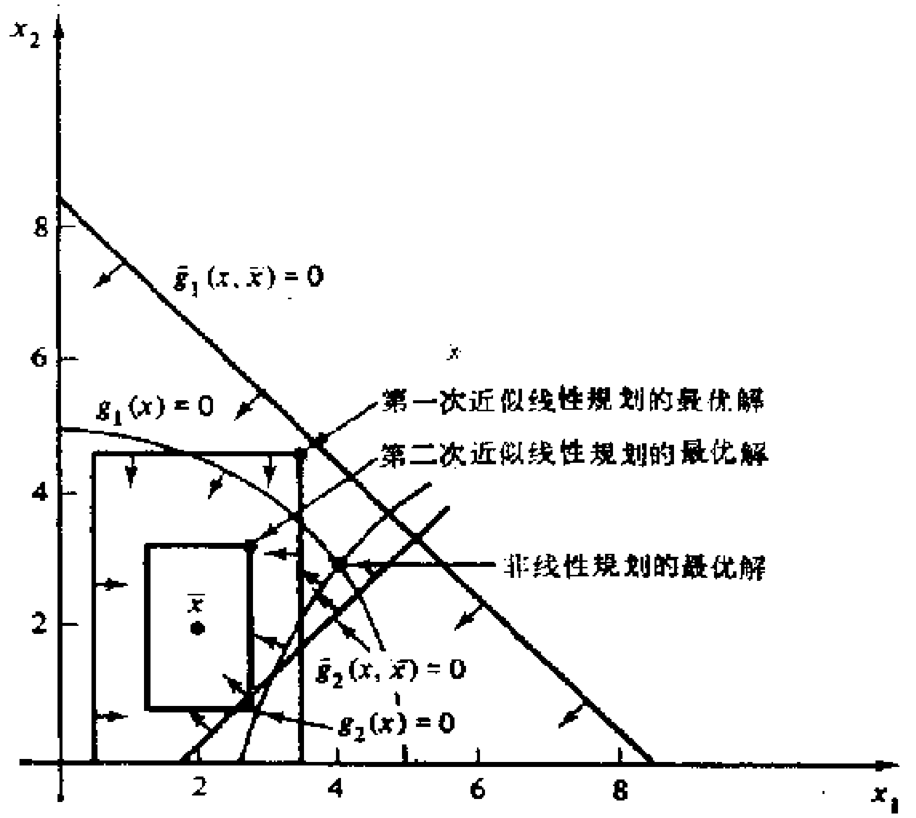


图 14.1 线性近似规划方法

$$2\geqslant y_2^-\geqslant 0. \qquad (14.29)$$

最优解是 $y_1^+=1.5$, $y_2^+=2.5$, $y_1^-=y_2^-=0$, 用原来的变量, 是 $x_{L,1}^*=3.5$, $x_{L,2}^*=4.5$. 这个点不满足(14.20), 所以必须缩小 m_j. 我们将它缩小一半, 取 $m_1=0.75$, $m_2=1.25$, 新的最优解在 $x_{L,1}^*=2.75$, $x_{L,2}^*=3.25$. 这个点对于原来的规划是能行的, 从而作为新的近似点. ▌

应该提一下, 函数是可分离的非线性规划, 例如目标函数为

$$f(x)=f_1(x_1)+f_2(x_2)+\cdots+f_n(x_n), \qquad (14.30)$$

且约束为线性的(或可分离的)规划, 用线性近似去解是很方便的, 其中使用了线性规划单纯形法的某种修改. 关于细节, 读者可以参考 Beale[7] 和 Dantzig[12].

成功地实现近似规划方法, 取决于近似子问题(如 Griffith-Stewart 方法中的线性规划)的有效求解. 有效的二次规划算法, 加上用二次函数能更精确地逼近非线性函数这一事实, 结果发展为更新近的近似规划方法. 这类算法的一个代表是下面描述的

Fletcher 的超方体法[17].

这个方法是为了求解线性约束或可能是无约束的问题而设计的. Fletcher 没有考虑非线性约束问题，也许因为这种约束能较好地间接处理，也就是说，把它们变换到目标函数中去，正如在第12章中看到的那样. 所以，考虑问题

$$\min f(x) \tag{14.31}$$

受限制于

$$A^T x \geqslant b, \tag{14.32}$$

其中 f 是可微实函数，A 是 $n\times m$ 矩阵，b 是 m 维向量. 令 X 表示满足 (14.32) 的 $x\in R^n$ 的集合. 在这个方法的每次迭代中，二次近似 $f_q(x, x^k)$ 在 x^k 点与 f 在 x^k 点取值相同，且

$$\nabla f_q(x^k)=\nabla f(x^k).$$

又用一个给定的对称矩阵 B_k 作为 f_q 的 Hesse 阵. 每一个子问题由求解一个二次规划组成，要在约束(14.32)下对 f_q 取极小，且从当前点 x^k 出发的移动要限制在一个超方体内，换言之，Chebyshev 模要满足

$$\|x-x^k\|_\infty=\max_{1\leqslant j\leqslant n}|x_j-x_j^k|\leqslant h^k. \tag{14.33}$$

下面将看到，这方法的一个重要特点是系统地调整 h^k，使得在(14.33)定义的超方体内，f_q 在某种意义下有效地近似于 f.

另一个重要的特点是每次迭代后矩阵 B_k 的修改. 如同在某些无约束最优化的拟-Newton 算法那样，B_k 用秩 2 校正来修改，以使下列形式的割线关系式满足：

$$B_k p^k=\gamma^k, \tag{14.34}$$

其中 $p^k=x^k-x^{k-1}$、$\gamma^k=\nabla f(x^k)-\nabla f(x^{k-1})$. 下面的修改公式是属于 Powell[27]:

$$B_k=B_{k-1}+\frac{(\gamma^k-B_{k-1}p^k)(p^k)^T+p^k(\gamma^k-B_{k-1}p^k)^T}{(p^k)^Tp^k}$$
$$-\frac{(p^k)^T(\gamma^k-B_{k-1}p^k)p^k(p^k)^T}{[(p^k)^Tp^k]^2}. \tag{14.35}$$

这个修改公式的某些性质和相应的无约束变尺度方法已在练习 11.L 中考察过.

如果 B_{k-1} 是正定的，修改公式(14.35)并不能保证 B_k 是正定的，结果是，二次子问题可能有满足 Kuhn-Tucker 最优性必要条件但不是整体极小的点. 虽然 Fletcher 的二次规划算法[16]能有效地处理上述情况，但更希望的也许是对修改公式作出限制，使得从正定的 B_0 出发，以后所有的 B_k 也是正定的. 这个步骤可以通过下述办法实现：(14.35)右端第二项和第三项乘以一个适当选择的常数 θ，$0\leqslant\theta\leqslant1$，使得[26]

$$\det B_k \geqslant \alpha \det B_{k-1}, \tag{14.36}$$

其中，比如说 $\alpha=0.1$. Fletcher 的超方体法，把可靠的理论基础同经验确定的常数结合起来，以保证成功的应用. 这方法现在可叙述如下：

1. 假设 $x^0\in X$ 是已给的，计算 $f(x^0)$、$\nabla f(x^0)$，且选择常数 $\bar{h}$、h^0，使得 $\bar{h}\geqslant h^0>0$. 取 $k=0$.

2. 如果 x^k (在某个允许范围内)满足下列 Kuhn-Tucker 条件，就停止：

$$\nabla f(x^k)-\sum_{i=1}^{m}\lambda_i a^i=0, \tag{14.37}$$

$$\lambda_i[(a^i)^T x^k-b_i]=0, \quad i=1, \cdots, m, \tag{14.38}$$

$$\lambda\geqslant 0. \tag{14.39}$$

否则，用(14.35)确定的 B_k，或对修改公式作出某种变动，以保证 B_k 是正定的，这里 B_0 是任意的对称(正定) $n\times n$ 阵.

3. 求解下列二次子问题：

(QSUB)

$$\min f_Q(x, x^k)=f(x^k)+(x-x^k)^T\nabla f(x^k)+\frac{1}{2}(x-x^k)^T B_k(x-x^k) \tag{14.40}$$

受限制于

$$A^T x\geqslant b, \tag{14.41}$$

$$\|x-x^k\|_\infty\leqslant h^k, \tag{14.42}$$

并设其解为 x^{k*}. 计算 $f(x^{k*})$、$\nabla f(x^{k*})$ 和

$$\rho^k=\frac{f(x^k)-f(x^{k*})}{f_Q(x^k, x^k)-f_Q(x^{k*}, x^k)}, \tag{14.43}$$

如果 $\rho^k \geqslant 0.5$，取 $x^{k+1}=x^{k*}$ 并转向步骤 4；否则转向步骤 5.

4. 如果下述条件中任何一个成立，那么取 $k=k+1$ 并转向步骤 2：

$$\|x^{k*}-x^k\|_\infty < h^k, \tag{14.44}$$

$$I(x^{k*}) \not\subset I(x^k), \tag{14.45}$$

其中 $I(x^{k*})$ 和 $I(x^k)$ 分别是在 x^{k*} 和 x^k 处 (14.41) 中积极约束指标集. 否则，求解一维极小化问题：

$$\min_{\alpha} f_q(x^k+\alpha(x^{k*}-x^k)) \tag{14.46}$$

受限制于

$$|\alpha| \leqslant 2, \tag{14.47}$$

且设 α^* 是其解. 计算 $\bar{x}=x^k+\alpha^*(x^{k*}-x^k)$，如果

$$f_q(\bar{x},\ x^k) \leqslant f(x^k) - \frac{4|f(x^{k*})-f_q(x^{k*},\ x^k)|}{1-\rho^k}, \tag{14.48}$$

那么，令

$$h^{k+1}=\min\{\bar{h},\ 2h^k\}; \tag{14.49}$$

否则令 $h^{k+1}=h^k$. 取 $k=k+1$ 并转向步骤 2.

5. 如果 $\rho^k<0.1$，将 h^k 的当前值减小一半并返回到步骤 3. 如果 $0.1 \leqslant \rho^k < 0.5$，取 $x^{k+1}=x^{k*}$，$h^{k+1}=\frac{1}{2}h^k$，$k=k+1$ 并转向步骤 2.

Fletcher 指导了某些数值试验，得到了有希望的结果，关于算法的细节以及某些收敛性结果，读者可参考[17]. 没有约束的超方体法的收敛性质，以及其他有关的算法，也已为 Powell[28] 所研究.

在结束本节时，我们提一下某些早期的方法，它们以类似于超方体法的想法作为基础，这些方法原来是为无约束问题导出的.

假设我们要解变量为 $x_1, \cdots, x_n$ 的一个非线性方程组

$$f_i(x)=0, \quad i=1, \cdots, m, \tag{14.50}$$

其中 $m>n$. 一种处理是使得最小二乘方意义下的偏差取极小值，即求解

$$\min_x f(x)=\sum_{i=1}^{m}[f_i(x)]^2. \tag{14.51}$$

我们能在点 x^k 的周围将每个 f_i 线性化，得到二次近似规划

$$\min_x \sum_{i=1}^{m} [f_i(x^k)+(x-x^k)^T\nabla f_i(x^k)]^2. \tag{14.52}$$

设 $J(x^k)$ 是各列为 $\nabla f_i(x^k)$ 的 $n\times m$ Jacobi 矩阵. 变量为 $p^k=x-x^k$ 的二次函数的逗留点，可以由求解下列线性方程组得到：

$$J(x^k)(J(x^k))^T p^k=-J(x^k)[f_1(x^k),\ \cdots,\ f_m(x^k)]^T. \tag{14.53}$$

迭代求解(14.53)的一个算法是 Gauss 方法. 这个方法经常发散，因为 f 的二次近似仅在近似点的小邻域中有效. Levenberg[21]和 Marquardt[23] 修正了 Gauss 方法，在 (14.52) 中增加一个“阻尼项”，并求解问题

$$\min_x \sum_{i=1}^{m} [f_i(x^k)+(x-x^k)^T\nabla f_i(x^k)]^2+\frac{1}{2}\lambda^k\|x-x^k\|^2, \tag{14.54}$$

其中 $\lambda^k\geqslant 0$ 是一个参数，从一次迭代到另一次迭代时须进行调整. 在某种意义上，这个问题等价于一个约束最优化问题，其中 (14.52) 是目标函数，约束 $\|x-x^k\|^2\leqslant h$ 是强加的. 在第 10 章中简短地叙述过的 Goldfeld、Quandt、Trotter 的方法[18]，也是一个类似的小步长二次近似规划方法. 最后，Wilson[36] 的方法是与超方体法密切有关的，不同之处在于矩阵 B_k 取为 f 的精确 Hesse 阵 $\nabla^2 f(x^k)$. Wilson 方法的讨论可以在 Beale[7] 中找到.

14.2 既约梯度法

本节的讨论集中于求解有约束的非线性规划，其方法类似于解线性规划的单纯形法. 首先考察线性约束的既约梯度法，然后将它推广到非线性约束问题. 考虑线性约束问题

(LEP)
$$\min f(x) \tag{14.55}$$

受限制于

$$Ax=b, \tag{14.56}$$

$$x\geqslant 0, \tag{14.57}$$

其中，f 是一个连续可微的实函数，A 是 $m\times n$ 阵，b 是 m 维向量

且 $m\leqslant n$. 变数向量 x 能划分为两个子向量 $x=(x^B, x^N)^T$, 其中 $x^B=(x_1^B, \cdots, x_m^B)^T$ 为基本变数向量或依从变数向量, $x^N=(x_1^N, \cdots, x_{n-m}^N)$ 为非基本变数向量, 或独立变数向量. 相应地, 矩阵 A 也划分为 $A=[B, C]$, 不失一般性, 这里假设矩阵 A 的前 m 列对应于基本变量. 进一步假设, 对应于向量 x^B 中分量的 A 的 $m\times m$ 子矩阵 B 是非异的. 于是可以写出

$$Bx^B+Cx^N=b, \tag{14.58}$$

从而

$$x^B=B^{-1}b-B^{-1}Cx^N. \tag{14.59}$$

此外, 我们还假设基本变量也是非退化的, 即 $x^B>0$. 非基本变量称为独立的, 是因为对它们规定某些数值后, 可得到(14.58)的唯一解.

既约梯度法的基本想法是用 (14.59) 消去 x^B (作为 x^N 的函数), 并考虑只对 x^N 来说的最优化问题. 这个想法被用于下列方法中: Wilde[33], Wilde、Beightler[34] 通过"约束导数"的概念导出的微分算法, Wolfe[37,38] 的既约梯度法、Zangwill[41]的凸-单纯形法.

从(14.59)得到用下列公式表示的**既约梯度** $r\in R^{n-m}$:

$$r(x^N)=\nabla_{x^N}f(x^B(x^N), x^N)-(B^{-1}C)^T\nabla_{x^B}f(x^B(x^N), x^N). \tag{14.60}$$

现在, 如果能从 x^N 的当前值沿负的既约梯度方向作一个微小的移动, 而不违反关于向量 x 的非负性约束, 那么 f 的函数值将减小. 这个步骤可实现如下: 给定一个能行点 x^k, 对 $i=1, \cdots, n-m$ 计算

$$z_i^{N,k+1}=\begin{cases}0, & x_i^{N,k}=0 \text{ 且 } r_i(x^{N,k})>0,\\ -r_i(x^{N,k}), & \text{其他}.\end{cases} \tag{14.61}$$

并令

$$z^{B,k+1}=-B^{-1}Cz^{N,k+1}, \tag{14.62}$$

于是 $z^{k+1}=(z^{B,k+1}, z^{N,k+1})^T$. 下一个点 $x^{k+1}=(x^{B,k+1}, x^{N,k+1})^T$ 为

$$x^{k+1}=x^k+\alpha_{k+1}^*z^{k+1}, \tag{14.63}$$

其中 α_{k+1}^* 由下列关系式算得:

$$\alpha_{k+1}^1=\max\{\alpha_{k+1}:x^{B,k}+\alpha_{k+1}z^{B,k+1}\geqslant 0\}, \tag{14.64}$$

$$\alpha_{k+1}^2=\max\{\alpha_{k+1}:x^{N,k}+\alpha_{k+1}z^{N,k+1}\geqslant 0\}, \tag{14.65}$$

并且

$$\begin{aligned} &f(x^k+\alpha_{k+1}^*z^{k+1})\\ &\quad=\min_{\alpha_{k+1}}\{f(x^k+\alpha_{k+1}z^{k+1}):0\leqslant\alpha_{k+1}\leqslant\min(\alpha_{k+1}^1,\ \alpha_{k+1}^2)\},\end{aligned} \tag{14.66}$$

当 $\alpha_{k+1}^*<\alpha_{k+1}^1$ 时, x^{k+1} 由(14.63)确定; 否则, 对某个 l 有

$$x_l^{B,k}+\alpha_{k+1}^1 z_l^{B,k+1}=0, \tag{14.67}$$

就从基本变量中除去 x_l^B, 而换入最大正值的非基本变量. 如果 $\|z^{k+1}\|\leqslant\varepsilon$, 算法便终结, 其中 $\varepsilon>0$ 是某个小的预定数值.

例 14.2.1

我们在如下问题上作几次迭代来说明既约梯度法:

$$\min f(x)=(x_1)^2+4(x_2)^2 \tag{14.68}$$

受限制于

$$x_1-2x_2-x_3\qquad=1, \tag{14.69}$$

$$-x_1+\ x_2\qquad+x_4=0, \tag{14.70}$$

$$x\geqslant 0. \tag{14.71}$$

假如从 $x^0=(2,\ 1,\ 3,\ 1)$ 开始, 且令 $x^{B,0}=(x_1,\ x_4)^T$, $x^{N,0}=(x_2,\ x_3)^T$. 于是

$$B_0=\begin{bmatrix}1&0\\-1&1\end{bmatrix},\quad (B_0)^{-1}=\begin{bmatrix}1&0\\1&1\end{bmatrix},\quad C_0=\begin{bmatrix}2&-1\\1&0\end{bmatrix}, \tag{14.72}$$

且 $\nabla f(x^0)=(4,\ 8,\ 0,\ 0)^T$. 我们计算既约梯度

$$r(x^{N,0})=\begin{pmatrix}8\\0\end{pmatrix}-\begin{bmatrix}2&1\\-1&0\end{bmatrix}\begin{bmatrix}1&1\\0&1\end{bmatrix}\begin{pmatrix}4\\0\end{pmatrix}=\begin{pmatrix}0\\4\end{pmatrix}. \tag{14.73}$$

因此 $z^{N,1}=(0,\ -4)^T$ 且

$$z^{B,1}=-\begin{bmatrix}1&0\\1&1\end{bmatrix}\begin{bmatrix}2&-1\\1&0\end{bmatrix}\begin{pmatrix}0\\-4\end{pmatrix}=\begin{pmatrix}-4\\-4\end{pmatrix}. \tag{14.74}$$

现在计算沿着 z^1 的步长．首先从 (14.64)求出 α_1^1．它是满足下式的最大的 α_1：

$$\begin{pmatrix} 2-4\alpha_1 \\ 1-4\alpha_1 \end{pmatrix} \geqslant 0, \tag{14.75}$$

从而得到 $\alpha_1^1=\dfrac{1}{4}$．类似地由下式求出 α_1^2：

$$\begin{pmatrix} 0+0\alpha_2 \\ 3-4\alpha_2 \end{pmatrix} \geqslant 0, \tag{14.76}$$

从而 $\alpha_1^2=\dfrac{3}{4}$．因此 $\alpha_1^1<\alpha_1^2$．关于 α_1 求 $f(x^0+\alpha_1 z^1)$ 的极小，从 (14.66) 得到 $\alpha_1^*=\dfrac{1}{4}$．于是，新的点为 $x^1=x^0+\alpha_1^* z^1$，即 $x^1=(1,1,2,0)$．在这个点上，$\nabla f(x^1)=(2,8,0,0)^T$．因为 $\alpha_1^*=\alpha_1^1$，我们改变基，x_3 进入基以代替 x_4．因此

$$B_1=\begin{bmatrix} 1 & -1 \\ -1 & 0 \end{bmatrix}, \quad B^{-1}=\begin{bmatrix} 0 & -1 \\ -1 & -1 \end{bmatrix}, \quad C_1=\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}. \tag{14.77}$$

新的既约梯度由下式给出：

$$r(x^{N,1})=\begin{pmatrix} 8 \\ 0 \end{pmatrix}-\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ -1 & -1 \end{bmatrix}\begin{pmatrix} 2 \\ 0 \end{pmatrix}=\begin{pmatrix} 10 \\ 2 \end{pmatrix}, \tag{14.78}$$

并由(14.61)得 $z^{N,2}=(-10,0)^T$．因此

$$z^{B,2}=-\begin{bmatrix} 0 & -1 \\ -1 & -1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}\begin{pmatrix} -10 \\ 0 \end{pmatrix}=\begin{pmatrix} -10 \\ -30 \end{pmatrix}, \tag{14.79}$$

其次我们计算沿着 z^2 的步长．不等式组

$$\begin{pmatrix} 1-10\alpha_2 \\ 2-30\alpha_2 \end{pmatrix} \geqslant 0 \tag{14.80}$$

和 $1-10\alpha_2\geqslant 0$ 必须成立，得到 $\alpha_2^1=\dfrac{1}{15}$，$\alpha_2^2=\dfrac{1}{10}$，且 $\alpha_2^1<\alpha_2^2$．关于 α_2 求 $f(x^1+\alpha_2 z^2)$ 的极小，所得到的没有受限制的步长是

$$\alpha_2=\frac{1}{10}>\alpha_2^1,$$

这样，我们取 $\alpha_2^* = \frac{1}{15}$.

下一个点由 $x^2 = \left(\frac{1}{3}, \frac{1}{3}, 0, 0\right)$ 给出，且

$$\nabla f(x^2) = \left(\frac{2}{3}, \frac{8}{3}, 0, 0\right)^T.$$

我们改变基，x_2 代替 x_3 变为一个基本变量. 现在 $x^{B,2} = (x_1, x_2)^T$，$x^{N,2} = (x_3, x_4)^T$. 矩阵 A 的新划分是

$$B_2 = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}, \quad (B_2)^{-1} = \begin{bmatrix} \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}, \quad C_2 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{14.81}$$

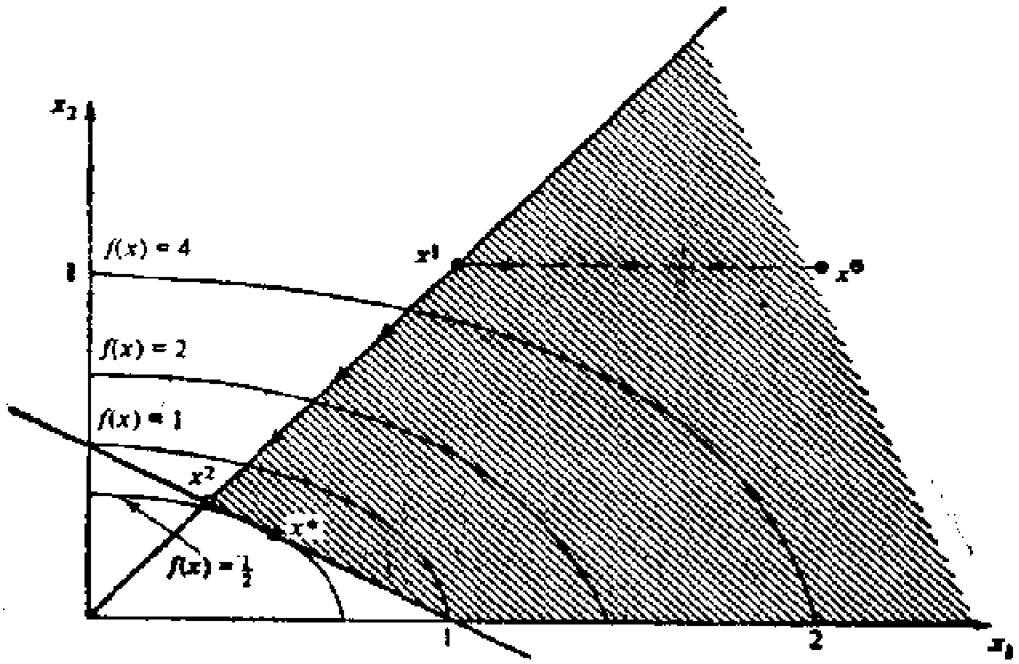


图 14.2 既约梯度法

在这点的既约梯度为

$$r(x^{N,2}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} \end{bmatrix} \begin{pmatrix} \frac{2}{3} \\ \frac{8}{3} \end{pmatrix} = \begin{pmatrix} \frac{10}{9} \\ -\frac{4}{9} \end{pmatrix}, \tag{14.82}$$

且 $z^{N,3}=\left(0,\ \frac{4}{9}\right)^T$. 这样

$$z^{B,3}=-\begin{bmatrix}\frac{1}{3} & -\frac{2}{3}\\ \frac{1}{3} & \frac{1}{3}\end{bmatrix}\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}\begin{pmatrix}0\\ \frac{4}{9}\end{pmatrix}=\begin{pmatrix}\frac{8}{27}\\ -\frac{4}{27}\end{pmatrix}. \qquad (14.83)$$

要求读者继续计算，并注意可收敛到最优点 $x^*=\left(\frac{1}{2},\ \frac{1}{4},\ 0,\ \frac{1}{4}\right)$. 这个算法的进程(在 x_1, x_2 空间)在图 14.2 中说明. ▌

上述既约梯度法由于锯齿现象而可能不收敛到满足 Kuhn-Tucker 条件的点. 已提出若干"反锯齿现象"的技巧去克服这个困难. 一个建议是 $z_i^{N,k+1}$ 的选择方法修改为：如果 $x_i^{N,k}\leqslant\varepsilon$ 且 $r_i(x^{N,k})>0$, 其中 $\varepsilon>0$ 是某个预先给定的值，则 $z_i^{N,k+1}$ 取为零，以此来代替(14.61)给出的公式. 另一个方法是 McCormick[24] 提出的：如果 α_k^* 不是 $f(x^{k-1}+\alpha_k z^k)$ 沿 z_k 的无约束极小，则 z^{k+1} 在 $\nabla f(x^{k-1})$ 的基础上进行计算.

如果 f 是线性函数，将基的改变规则作某些简易的修改，既约梯度法将成为线性规划的单纯形法. 如果没有约束出现，既约梯度法就成为最速下降法，这样，它可以看作是无约束方法的扩充. McCormick[25] 提出了既约梯度法的一个修改，它可以看作是线性约束问题的 DFP 变尺度算法的某种扩充. 他证明了，倘若某些温和的条件被满足，则他的方法以超线性速率收敛于 Kuhn-Tucker 点.

Abadie-Carpentier[1] 推广既约梯度法于具有非线性等式约束的问题. 考虑有界变量问题：

(NEP)
$$\min f(x) \qquad (14.84)$$

受限制于

$$h_i(x)=0,\quad i=1,\ \cdots,\ m, \qquad (14.85)$$

$$\beta\geqslant x\geqslant\alpha. \qquad (14.86)$$

其中 f, h_1, h_2, $\cdots$, h_m 假定是 $x\in R^n$ 的连续可微函数. 且 $m\leqslant n$.

每个能行向量 x^0 假定是非退化的，与前面一样，这意味着 x^0

可以划分为两个子向量 $x=(x^{B,0}, x^{N,0})^T$，其中

$$x^{B,0}\in R^m, \quad x^{N,0}\in R^{n-m}.$$

类似地，$\alpha=(\alpha^{B,0}, \alpha^{N,0})^T$，$\beta=(\beta^{B,0}, \beta^{N,0})^T$，使得

$$\beta^{B,0}>x^{B,0}>\alpha^{B,0}, \tag{14.87}$$

且向量

$$\nabla_{x^B}h_i(x^0)=\left[\frac{\partial h_i(x^0)}{\partial x_1^{B,0}}, \cdots, \frac{\partial h_i(x^0)}{\partial x_m^{B,0}}\right]^T, \quad i=1, \cdots, m \tag{14.88}$$

是线性独立的. 也就是说，各列为向量 $\nabla_{x^B}h_i(x^0)$ 的 $m\times m$ 阵 $\Delta^B(x^0)$ 是非异的. 容易验证，在 x^0 处问题(NEP)最优性的 Kuhn-Tucker 必要条件是，存在向量 $\mu^0\in R^m$, $\lambda^0\in R^{n-m}$，使得

$$\nabla_{x^N}f(x^0)-\sum_{i=1}^{m}\mu_i^0\nabla_{x^N}h_i(x^0)-\lambda^0=0, \tag{14.89}$$

$$\nabla_{x^B}f(x^0)-\sum_{i=1}^{m}\mu_i^0\nabla_{x^B}h_i(x^0)=0, \tag{14.90}$$

其中

$$\lambda_j^0\geqslant 0, \ 若\ x_j^{N,0}=\alpha_j^{N,0}, \tag{14.91}$$

$$\lambda_j^0=0, \ 若\ \beta_j^{N,0}>x_j^{N,0}>\alpha_j^{N,0}, \tag{14.92}$$

$$\lambda_j^0\leqslant 0, \ 若\ x_j^{N,0}=\beta_j^{N,0}, \tag{14.93}$$

因此

$$\mu^0=[\Delta^B(x^0)]^{-1}\nabla_{x^B}f(x^0), \tag{14.94}$$

$$\lambda^0=\nabla_{x^N}f(x^0)-\Delta^N(x^0)[\Delta^B(x^0)]^{-1}\nabla_{x^B}f(x^0), \tag{14.95}$$

其中 $\Delta^N(x^0)$ 是 $(n-m)\times m$ 阵，它的列是向量 $\nabla_{x^N}h_i(x^0)$. 注意，如果约束是线性的，λ^0 就是类似于(14.60)的既约梯度. 由此可知，对于每个能行的 x^0，我们可以利用上一方程计算向量 λ^0；如果 λ^0 还满足(14.91)至(14.93)，那么(NEP)的最优性必要条件在 x^0 处成立. **广义既约梯度法**基于从一个能行的 x^0 以迭代方式移动到能行的 x^1, x^2, 等等，直到抵达点 x^l，在 x^l 处由(14.95)算出的 λ^l 满足(14.91)至(14.93). 假设有一个能行的 x^k 而相应的 λ^k 不满足这三个关系式，那就按下式改变非基本向量的当前值为

$$x^{N,k+1}=x^{N,k}+\theta_{k+1}z^{k+1}, \quad \theta_{k+1}\geqslant 0, \tag{14.96}$$

其中 $z_j^{k+1}=0$, 如果

$$x_j^{N,k}=\alpha_j^{N,k} \text{ 和 } \lambda_j^k>0, \quad \text{或 } x_j^{N,k}=\beta_j^{N,k} \text{ 和 } \lambda_j^k<0. \qquad (14.97)$$

否则，按照依赖于这方法的具体做法的规则，取 $z_j^{k+1}=-\lambda_j^k$ 或 $z_j^{k+1}=0$.

1．在 GRG 做法中，每当(14.97)不成立时我们取 $z_j^{k+1}=-\lambda_j^k$. 注意，如同我们上面看到的，如果约束是线性的，那末 z_j^{k+1} 的这种选法和 Wolfe[38] 所用的选法是一致的.

2．在 GRGS 做法中，首先用下式求出指标 s:

$$|\lambda_s^k|=\max_j|\lambda_j^k|, \qquad (14.98)$$

其中指标 j 在(14.97)不成立的指标集上变化．然后令

$$z_j^{k+1}=\begin{cases}0, & j\neq s,\\ -\lambda_k^s, & j=s.\end{cases} \qquad (14.99)$$

如果最优化问题实际上是一个线性规划，那么这个做法与单纯形法一致.

3．在 GRGC 的做法中，循环地取 $z_j^{k+1}=-\lambda_j^k$. 一个循环由 n 次迭代组成．对于 $k=1, \cdots, n$, 我们取

$$z_j^{k+1}=\begin{cases}0, & j\neq k,\\ -\lambda_j^k, & j=k,\end{cases} \qquad (14.100)$$

除非 k 是一基本变量的指标或(14.97)成立，这时第 k 次迭代省略．在 $k=n$ 之后转回到 $k=1$, 如此等等.

用这些做法之一确定 $z^{k+1}\in R^{n-m}$ 后，取步长 θ_{k+1} 的一个试探值，比如 $\theta_{k+1}^1>0$, 并确定

$$\tilde{x}_j^{N,k}=\begin{cases}\beta_j^{N,k}, & x_j^{N,k}+\theta_{k+1}^1 z_j^{k+1}>\beta_j^{N,k},\\ \alpha_j^{N,k}, & x_j^{N,k}+\theta_{k+1}^1 z_j^{k+1}<\alpha_j^{N,k},\\ x_j^{N,k}+\theta_{k+1}^1 z_j^{k+1}, & \text{其他}.\end{cases} \qquad (14.101)$$

为了保持能行性，用 Newton 法之类的迭代过程，解 m 个未知量 $x_1^B, \cdots, x_m^B$ 的 m 个非线性方程:

$$h_i(x^B, \tilde{x}^{N,k})=0, \quad i=1, \cdots, m. \qquad (14.102)$$

于是，在 Newton 法的第 l 次迭代，有公式

$$x^{B,l+1}=x^{B,l}-[\Delta^N(x^{B,l},\ \tilde{x}^{N,k})]^{-1}h(x^{B,l},\ \tilde{x}^{N,k}),\qquad(14.103)$$

其中 $h=(h_1,\ \cdots,\ h_m)^T$，且对 $l=0$ 有 $x^{B,l}=x^{B,k}$. 公式(14.103)一直使用到下述情况之一发生：

1．在几个相继的迭代中，模 $\|h(x^{B,l},\ \tilde{x}^{N,k})\|$ 增大．这时，用某个因子将 θ_{k+1}^1 减小，计算(14.101)，然后返回求解(14.102).

2．我们得到

$$f(x^{B,l},\ \tilde{x}^{N,k})>f(x^{B,l-1},\ \tilde{x}^{N,k}),\qquad(14.104)$$

这时的校正与情况 1 相同，

3．对某个 l，点 $(x^{B,l},\ \tilde{x}^{N,k})$ 在(14.86)确定的范围之外．这时，在连接 $x^{B,l-1}$ 与 $x^{B,l}$ 的线段上求一点 $\tilde{x}^{B,l}$，使得对某个 r 有 $\tilde{x}_r^{B,l}=\alpha_r^{B,l}$ 或 $\tilde{x}_r^{B,l}=\beta_r^{B,l}$，并改变当前的基，比如说以 $x_s^{N,l}$ 代替变量 $x_r^{B,l}$，其中指标 s 用类似于单纯形法中的某个规则来确定．例如，可以用关系式(略去了迭代指标)：

$$|\Delta^N(\tilde{x})_s(\Delta^B(\tilde{x}))_r^{-1}|v_s=\max_p\{|\Delta^N(\tilde{x})_p(\Delta^B(\tilde{x}))_r^{-1}|v_p\},\qquad(14.105)$$

其中 $\Delta^N(\tilde{x})_p$ 是行向量 $(\partial h_1(\tilde{x})/\partial x_p^N,\ \cdots,\ \partial h_m(\tilde{x})/\partial x_p^N)$ 在 $\tilde{x}=(\tilde{x}^{B,l},\ \tilde{x}^{N,k})$ 的值，$(\Delta^B(\tilde{x}))_r^{-1}$ 是矩阵 $(\Delta^B(\tilde{x}))^{-1}$ 的第 r 列，而

$$v_p=\min\{(\tilde{x}_p^{N,k}-\alpha_p^{N,k}),\ (\beta_p^{N,k}-\tilde{x}_p^{N,k})\}.\qquad(14.106)$$

用新的基试解(14.102).

4．迭代收敛，即对某个 l^*，

$$\|h(x^{B,l^*},\ \tilde{x}^{N,k})\|\leqslant\varepsilon,\qquad(14.107)$$

其中 ε 是一个小的正数．若

$$\beta^{B,l^*}>x^{B,l^*}>\alpha^{B,l^*},\qquad(14.108)$$

则取 $x^{k+1}=(x^{B,l^*},\ \tilde{x}^{N,k})$，并求一个新的搜索方向．若 x^{B,l^*} 的第 r 个分量恰是上界或下界，则按情形 3 一样改变基．当满足最优性必要条件的一点找到时，计算结束．

进行计算的实际方案包含很多重要的考虑，读者无疑已认识到，既约梯度法用于非线性约束是不简单的．有趣的是，在

Colville[11] 关于非线性规划程序的比较性研究中，广义既约梯度法在试验的方法中是最好的一个.

14.3 割平面法

到目前为止，我们研究了凸性及其推广不起主要作用的计算方法. 尽管如此,很多算法对于凸规划可以更容易地实现,这时迭代的极限点通常满足最优性的必要充分条件，这是一个并非每个非线性规划都具有的特性. 这里讨论的方法适用于只有某种凸性的非线性规划. 割平面法的基本原理是，非线性规划的能行集用有限个闭半空间来近似,并求解一系列近似的线性规划.

考虑问题

(I) $$\min \hat{g}_0(x) \tag{14.109}$$

受限制于

$$g_i(x) \geqslant 0, \quad i=1, \cdots, l. \tag{14.110}$$

其中 $\hat{g}_0$ 是闭正常凸函数，g_i 是闭正常凹函数，$i=1, \cdots, l$. 这个非线性规划容易修改成具有线性目标函数的等价规划. 定义一个新的变量 $x_0 \in R$，我们写出

(II) $$\min x_0 \tag{14.111}$$

受限制于

$$g_i(x) \geqslant 0, \quad i=0, 1, \cdots, l, \tag{14.112}$$

其中

$$g_0(x) = x_0 - \hat{g}_0(x) \tag{14.113}$$

是闭正常凹函数. 这样，通过增加一个附加变量和一个附加约束，我们将规划 (I) 变为 (II)，使得 x^* 为 (I) 的解的充要条件是 (x_0^*, x^*) 为 (II) 的解，且 $x_0^* = \hat{g}_0(x^*)$. 所以，不失一般性，可以考虑凸规划问题

(CPP) $$\min f(x) = c^T x \tag{14.114}$$

受限制于

$$g_i(x) \geqslant 0, \quad i=1, \cdots, m, \tag{14.115}$$

其中 $x=(x_1, x_2, \cdots, x_n)^T$，$g_1, \cdots, g_m$ 是 R^n 上的闭正常凹函数.

令

$$G=\{x: g_i(x)\geqslant 0,\ i=1,\ \cdots,\ m\}, \tag{14.116}$$

并假定 G 是包含在由有限个线性不等式定义的一个紧致集 $T\subset R^n$ 中，这里 T 为

$$T=\{x: x\in R^n,\ Ax\geqslant b\}. \tag{14.117}$$

对于 $x\in T$，令 $\partial g_i(x)\subset R^n$ 是 g_i 的次梯度的集合，或次微分(参见 4.3 节)，即 $\xi\in\partial g_i(x)$ 蕴涵

$$g_i(y)\leqslant g_i(x)+\xi^T(y-x) \tag{14.118}$$

对于每个 $y\in R^n$ 成立. 如果 g_i 在 x 是可微的，那么 $\xi=\nabla g_i(x)$. 我们将假定，对于 $x\in T$，

$$\partial g_i(x)\neq\emptyset,\quad i=1,\ \cdots,\ m, \tag{14.119}$$

并且集合

$$\bigcup_{x\in T}\partial g_i(x),\quad i=1,\ \cdots,\ m \tag{14.120}$$

是有界的. 例如，当 g_i 连续可微时，这个假定是满足的. 我们现在可以叙述求解 (CPP) 的**割平面算法**，它是 Kelley[20] 和 Cheney-Goldstein[10] 导出的，我们将称之为 KCG 算法:

1. 求解线性规划: $\min f(x)=c^Tx$，受限制于 $x\in T$，且设 x^0 是最优解. 若 x^0 属于集合

$$G(\varepsilon)=\{x: x\in T,\ g_i(x)\geqslant-\varepsilon,\ i=1,\ \cdots,\ m\}, \tag{14.121}$$

其中 ε 是一个小的正数，则算法停止; (CPP) 的一个最优解就达到了. 否则，令 $k=0$ 并进到步骤 2.

2. 给定了 $x^k\in T$，且 $x^k\notin G(\varepsilon)$ 时，根据

$$g_{s_k}(x^k)=\min\{g_i(x^k),\ i=1,\ \cdots,\ m\}<0 \tag{14.122}$$

求出指标 s_k，选一个 $\xi^k\in\partial g_{s_k}(x^k)$. 并求解线性规划:

$$\min c^Tx, \tag{14.123}$$

$$\tilde{g}_{s_h}(x,\ x^h)=g_{s_h}(x^h)+(\xi^h)^T(x-x^h)\geqslant 0,\ h=0,\ 1,\ \cdots,\ k, \tag{14.124}$$

$$x\in T. \tag{14.125}$$

3. 令 x^{k+1} 是上述线性规划的最优解. 若 $x^{k+1}\in G(\varepsilon)$，则算

法停止，否则令 $k=k+1$ 并转回步骤 2.

注意，由 g_i 的凹性推得的 (14.118) 和约束 (14.124) 解释了割平面这名称. 假设 $x^k \overline{\in} G(\varepsilon)$，于是至少有某个指标 s 使得 $g_s(x^k) < -\varepsilon$. 设 $\xi \in \partial g_s(x^k)$，由

$$g_s(x^k) + \xi^T(x - x^k) \geqslant 0 \tag{14.126}$$

定义的半空间从能行集 G 中割去了不满足(14.126)的点 x^k，而 G 包含在 (14.126) 定义的半空间中. 用 S_k 表示第 k 次迭代的第 2 步中被求解的线性规划的能行集. 这些集合是一个套一个的，也就是

$$S_k \subset S_{k-1} \subset \cdots \subset S_0. \tag{14.127}$$

在每次迭代时，一个新的线性不等式约束加到前一次迭代时已有的约束集合中去. 在每次迭代时增加一个新约束，这使人们想到去利用求解线性规划问题的对偶单纯形法[12]. 现在叙述并证明凸规划 KCG 方法的收敛性.

定理 14.1

设 $g_1, \cdots, g_m$ 是紧致凸集 $T \subset R^n$ 上的闭凹函数，在每个点 $x \in T$ 上，对 $i=1, \cdots, m$，次梯度集 $\partial g_i(x)$ 是非空的，且存在 K 使得

$$\sup\{\|\xi^i\| : \xi^i \in \partial g_i(x),\ i=1, \cdots, m;\ x \in T\} \leqslant K. \tag{14.128}$$

又假定(CPP)的能行集 G 是非空的，且包含在 T 中. 令

$$S_k = S_{k-1} \cap \{x : \tilde{g}_{s_k}(x, x^k) \geqslant 0\}, \tag{14.129}$$

其中 $S_0 = T$. 如果 $x^{k+1} \in S_k$ 使得

$$f(x^{k+1}) = c^T x^{k+1} = \min\{c^T x : x \in S_k\}, \tag{14.130}$$

那么序列 $\{x^k\}$ 包含着一个子序列，它收敛于(CPP)的最优解.

【证明】 首先由(14.127)注意到 $\{f(x^k)\}$ 是单调递增的. 因此，如果 $\{x^k\}$ 包含一个子序列收敛于一点 $x^* \in G$，那么 $\{f(x^k)\}$ 收敛于 $f(x^*)$，且 x^* 是(CPP)的解. 现在假设 $\{x^k\}$ 没有一个子序列收敛于 G 中的点. 那么存在 $\alpha > 0$，使得对于 $h=0, 1, \cdots, k$ 有

$$g_{s_h}(x^h) = \min_i\{g_i(x^h),\ i=1, \cdots, m\} \leqslant -\alpha. \tag{14.131}$$

如果 x^{k+1} 使 c^Tx 在 S_k 上取极小值，那么 $x^{k+1}\in T$，且

$$g_{s_h}(x^h)+(\xi^h)^T(x^{k+1}-x^h)\geqslant 0,\quad h=0,\ 1,\ \cdots,\ k,\tag{14.132}$$

从上面的两个关系式和 Schwarz 不等式，可推得

$$\alpha\leqslant -g_{s_h}(x^h)\leqslant(\xi^h)^T(x^{k+1}-x^h)\leqslant K\|x^{k+1}-x^h\|.\tag{14.133}$$

因此对于指标的每个子序列 $\{k_p\}$，我们有

$$\|x^{k_p}-x^{k_q}\|\geqslant\frac{\alpha}{K},\quad q<p.\tag{14.134}$$

即 $\{x^k\}$ 没有 Cauchy 子序列，但 $\{x^k\}\subset T$，且 $\{x^k\}$ 是有界的，这就得到矛盾. ▌

逐次构造割平面，并求解近似线性规划以找到点列 $\{x^k\}$，所得点列的典型行程示于图 14.3 中.

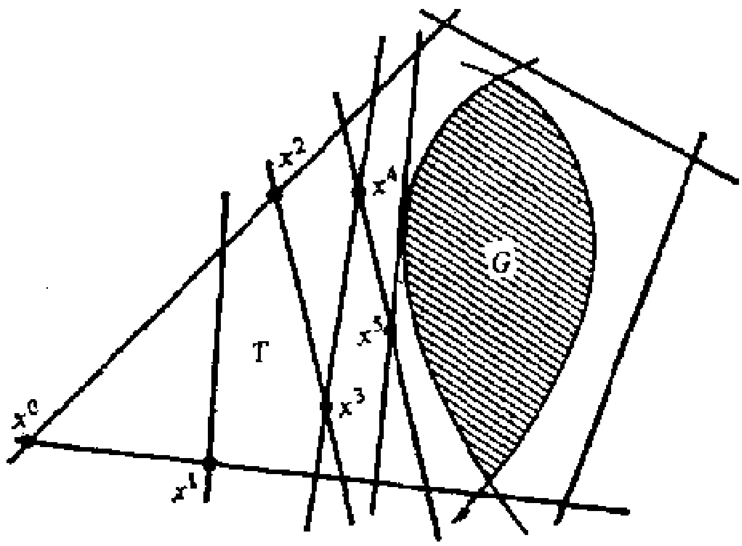


图 14.3 KCG 割平面算法

例 14.3.1

考虑凸规划问题：

$$\min f(x)=4x_1+5x_2\tag{14.135}$$

受限制于

$$g_1(x)=-(x_1)^2-2x_1x_2-2(x_2)^2+4\geqslant 0,\tag{14.136}$$

$$g_2(x)=-(x_1)^2-(x_2)^2+4x_1-3\geqslant 0.\tag{14.137}$$

令

$$T=\{x:x\in R^2,\ 4\geqslant x_1\geqslant 0,\ 4\geqslant x_2\geqslant -4\}.\tag{14.138}$$

应用 KCG 割平面算法，首先求解线性规划 $\min f(x)$，受限制于 $x\in T$. 最优解在 $x^0=(0,\ -4)$，在这点处 $f(x^0)=-20$，$g_1(x^0)=-28$，$g_2(x^0)=-19$，且 $s_0=1$. 因为约束 (14.136) 在 x^0 违反最多，我们构造线性约束

$$g_1(x^0)+(x-x^0)^T\nabla g_1(x^0)\geqslant 0, \tag{14.139}$$

即

$$2x_1+4x_2+9\geqslant 0. \tag{14.140}$$

求解线性规划：

$$\min f(x)=4x_1+5x_2 \tag{14.141}$$

受限制于

$$2x_1+4x_2+9\geqslant 0, \tag{14.142}$$

$$4\geqslant x_1\geqslant 0, \tag{14.143}$$

$$4\geqslant x_2\geqslant -4. \tag{14.144}$$

我们求得最优解在 $x^1=\left(\frac{7}{2},\ -4\right)$，在这点处

$$f(x^1)=-16,\quad g_1(x^1)=-\frac{49}{4},\quad g_2(x^1)=-\frac{69}{4}$$

且 $s_1=2$. 这样，我们增添一个约束

$$g_2(x^1)+(x-x^1)^T\nabla g_2(x^1)\geqslant 0, \tag{14.145}$$

即

$$-12x_1+32x_2+101\geqslant 0, \tag{14.146}$$

把它加到线性规划(14.141)至(14.144)中，求得新的最优解在

$$x^2=\left(\frac{29}{28},\ -\frac{155}{56}\right),$$

在这点处 $f(x^0)=-9.7$. ∎

要求读者绘制上述问题的图，并观察这个算法相当缓慢的进程. 对于某种特殊类型的凸规划，Levitin、Polyak[22] 和 Wolfe[39] 研究了 KCG 割平面法的收敛速率，得到的结果令人失望，因为发现一般说来不能期望有比线性更快的收敛速率. 尽管如此，与效率高的对偶单纯形法相结合后，对于求解适当规模的凸规划，割平面法仍是一种有用的方法. 在非凸规划的情况下，这

个方法完全无效，这是因为在每次迭代中加入的线性约束可能割去能行集中包含最优值点的那部分.

有趣的是，在割平面法的每次迭代中所要解的线性规划的对偶，其本身就是用 Dantzig、Wolfe[12,39] 的"广义规划"(分解)算法所要解的线性子问题.

也许影响到刚才所述方法的效率的最重要步骤是割平面的选法. 已对基本 KCG 方法提出了若干修改. Veinott[32] 把割平面法扩充到有拟凹约束函数的问题，如我们在第 6 章看到的，它确定了一个凸能行集. 割平面是用实际求出能行集的支撑超平面来构造的. 因此，Veinott 算法称为**支撑超平面法**，并被用来求解由(14.114)和(14.115)给出的问题 (CPP)，其中 $g_1, \cdots, g_m$ 假定为连续可微的拟凹函数. 由(14.116)确定的集合 G 是凸的，并假定它包含在由(14.117)确定的紧致集 T 中. 还假定存在一点 y，使得

$$g_i(y)>0, \quad i=1, \cdots, m. \tag{14.147}$$

这个方法由下列步骤组成：

1. 支撑超平面法的第 1 步与 KCG 算法的相应步骤一样.

2. 给定 $x^k\in T$，使 $x^k\notin G(\varepsilon)$，且 y 满足(14.147)，求一个数 θ^k，使得点

$$z^k=(1-\theta^k)y+\theta^k x^k \tag{14.148}$$

是能行的，且对某个 s_k 满足

$$g_{s_k}(z^k)=0. \tag{14.149}$$

求解线性规划

$$\min c^T x, \tag{14.150}$$

$$(x-z^h)\nabla g_{s_h}(z^h)\geqslant 0, \quad h=0, 1, \cdots, k, \tag{14.151}$$

$$x\in T. \tag{14.152}$$

3. 设 x^{k+1} 是前一步骤中线性规划的最优解. 如果

$$x^{k+1}\in G(\varepsilon),$$

那么算法停止；否则置 $k=k+1$ 并转回步骤 2.

在图 14.4 中说明一个典型情况. 注意，虽然为了求得能行集

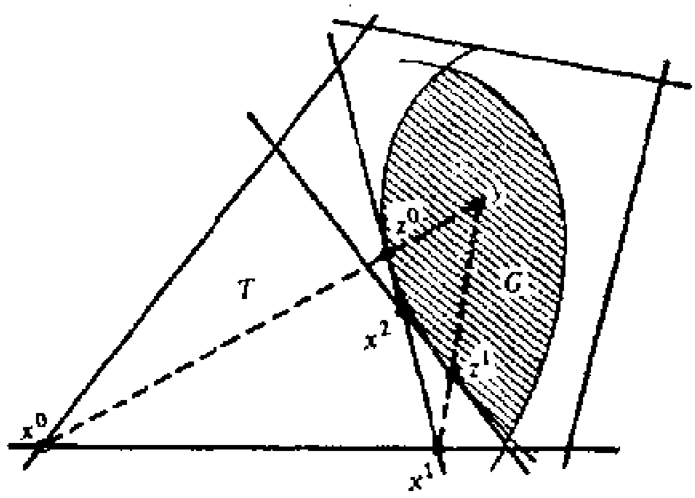


图 14.4 支撑超平面法

边界上的点，需要额外的计算，但是这使我们能构造出比 KCG 方法更深入的切割．支撑超平面法的收敛性可以在上面叙述过的假设下得到证明．

至今讨论过的割平面法的一个明显缺点是，从一次迭代到另一次迭代，线性规划子问题的规模不断增大，这是因为割平面约束总是加入已有的约束集中而从不删去．关于割平面法的某些稍晚的工作集中在去掉非积极约束的问题上，或更一般地，用线性约束逼近非线性规划的能行集的问题，而不要求一个套一个．Topkis[31] 导出了这样的方法，且在类似于本节早先所述的假定下证明了收敛性．Eaves、Zangwill[15] 在更一般的框架下研究了割平面法，它包括了割平面法的 KCG 做法，Veinott 做法，Topkis 做法，以及若干其他做法作为特殊情况．

14.4 相补凸规划

在本节中我们寻找线性函数在 R^n 中一个集合上的极小值点，这个集合是某些凸集和凸集的补集之交．这样的问题通常是一个非凸规划，称为**相补凸规划**或“反向”凸规划．Rosen[26] 和 Avriel、Williams[4] 导出的解法包含：用线性约束逐次逼近“反向”约束，从而得到一系列凸规划，用早先讨论过的方法之一，诸如惩罚函数法或割平面法，解这些凸规划．为了证明下述方法的收

敛性，我们依靠 Zangwill[40,41] 的一个相当一般的收敛性理论，它还能用来证明本书介绍的很多其他算法的收敛性.

考虑非线性规划问题

(CCP) $$\min f(x)=c^T x \tag{14.153}$$

受限制于

$$g_i(x)\geqslant 0,\quad i=1,\ \cdots,\ m, \tag{14.154}$$

$$h_l(x)\geqslant 0,\quad l=1,\ \cdots,\ p, \tag{14.155}$$

其中 $c\in R^n$ 是给定的非零向量，g_i 和 h_l 分别是实值连续可微凸函数和凹函数. 定义

$$T=\{x:x\in R^n,\ g_i(x)\geqslant 0,\ i=1,\ \cdots,\ m\}, \tag{14.156}$$

$$U=\{x:x\in R^n,\ h_l(x)\geqslant 0,\ l=1,\ \cdots,\ p\}, \tag{14.157}$$

并令 $X=T\cap U$. 注意，因为当 g_i 非线性时，集合 T 非凸，所以 (CCP) 一般不是凸规划. 实际上，T 是 m 个开凸集的补集之交. 我们特别感兴趣于 g_i 是凸的和非线性的情况. 因为我们处理的是非凸规划，所以不能保证算法收敛于 f 的整体约束极小值点，但是我们将证明，除了某些退化的情况之外，算法收敛到局部极小值点. 对于(CCP)，假定成立下述正则性条件：

1. 能行集 X 是紧致的且规划(CCP)是强相容的. 也就是至少存在一点 $x\in X$，使得(14.154)和(14.155)作为严格不等式成立.

2. 对于 $\tilde{x}\in X$，定义积极约束集为

$$I(\tilde{x})=\{i:g_i(\tilde{x})=0\}, \tag{14.158}$$

$$L(\tilde{x})=\{l:h_l(\tilde{x})=0\}. \tag{14.159}$$

于是，对于所有的数 $r_i,\ i\in I(\tilde{x})$ 和 $q_l,\ l\in L(\tilde{x})$，只要成立

$$r_i\geqslant 0,\quad q_l\geqslant 0,\quad \sum_{i\in I(\tilde{x})} r_i+\sum_{l\in L(\tilde{x})} q_l>0, \tag{14.160}$$

便对某个 $j,\ j=1,\ \cdots,\ n$，有

$$\sum_{i\in I(\tilde{x})} r_i\frac{\partial g_i(\tilde{x})}{\partial x_j}+\sum_{l\in L(\tilde{x})} q_l\frac{\partial h_l(\tilde{x})}{\partial x_j}\neq 0. \tag{14.161}$$

3. 令 ϕ 是 $x\in R^n$ 的实函数，如果存在一个向量 $\hat{x}$ 使得 $\phi(\hat{x})<0$，就称不等式 $\phi(x)\geqslant 0$ 是非平凡的. 假定在(14.154)和(14.155)

中的每一个不等式都是非平凡的.

满足最优性 Kuhn-Tucker 必要条件(3.72)至(3.74)的能行点叫做 **K-T** 点. 可以证明, 上述正则性条件蕴涵着, 对 (CCP) 的每个解存在 Kuhn-Tucker 乘子. 此外, 每个 K-T 点 x^* 必定满足

$$I(x^*)\cup L(x^*)\neq\emptyset. \tag{14.162}$$

后一个结果由(3.72)和(3.73)导出, 这是因为如果上面的集合是空的, 那么, 分别相应于约束(14.154)和(14.155)的所有 Kuhn-Tucker 乘子 λ_i^*, $i=1, \cdots, m$ 和 π_l^*, $l=1, \cdots, p$, 必定都为零, 这与 $\nabla f(x^*)=c$ 是一个非零向量相矛盾.

还可以验证, (CCP) 的局部 (或整体) 极大值点不能是 K-T 点. 因为, 假设 x^* 是 K-T 点, 那么它必须满足(14.162), 因此对于某个 j,

$$c_j=\sum_{i=1}^{m}\lambda_i^*\frac{\partial g_i(x^*)}{\partial x_j}+\sum_{l=1}^{p}\pi_l^*\frac{\partial h_l(x^*)}{\partial x_j}\neq 0, \tag{14.163}$$

从而, c^Tx 在 X 上取极大值的必要条件不能在 x^* 成立.

解问题 (CCP) 的算法叙述如下. 从某个能行点 $x^0\in X$ 出发, 按下面方式产生能行点序列 $\{x^k\}$: 给定一个点 $x^k\in X$, 对(14.154)中的函数 g_i, 用它关于 x^k 的一阶 Taylor 近似式来代替. 换言之, 用下列线性约束代替 $g_i(x)\geqslant 0$:

$$\tilde{g}_i(x, x^k)=g_i(x^k)+(x-x^k)^T\nabla g_i(x^k)\geqslant 0, \quad i=1, \cdots, m. \tag{14.164}$$

如果令

$$T(x^k)=\{x:\tilde{g}_i(x, x^k)\geqslant 0,\ i=1, \cdots, m\}, \tag{14.165}$$

我们得到一个新的规划:

$$(\mathrm{CCP}_k) \qquad \min c^Tx \tag{14.166}$$

受限制于

$$x\in X(x^k)=T(x^k)\cap U, \tag{14.167}$$

这是一个凸规划. 下一个点 x^{k+1} 取 (CCP_k) 的任一最优解. 这也就是, 当用 A^k 表示 (CCP_k) 的最优解的集合时, 我们从 A^k 中选取 x^{k+1}. 我们将证明, 如果 x^0 对 (CCP) 是能行的, 那么序列 $\{x^k\}$ 中

每一个也是这样．此外，还能证明，$\{x^k\}$ 的任何收敛子序列收敛于 (CCP) 的 K-T 点．

在转向收敛性问题之前，我们用一个简单的例子说明所提出的算法．

例 14.4.1

考虑规划

$$\min f(x)=x_1 \tag{14.168}$$

受限制于

$$g_1(x)=(x_1-3)^2+(x_2-2)^2-13\geqslant 0, \tag{14.169}$$

$$h_1(x)=-(x_1-4)^2-(x_2)^2+16\geqslant 0. \tag{14.170}$$

图 14.5 中阴影部分表示能行集，它显然是非凸的．在这个问题中有三个 K-T 点．两个局部极小值点分别在 (0, 0) 和(6.4, 3.2)，不是局部极小值点的 K-T 点在 $(3+\sqrt{13},\ 2)$．假设我们从 $x^0=(7,\ 0)$ 开始，然后用 g_1 关于 x^0 的一阶 Taylor 近似式代替 g_1:

$$\tilde{g}_1(x,\ x^0)=8x_1-4x_2-49, \tag{14.171}$$

并求解下列凸规划：

(CCP$_0$) $$\min f(x)=x_1 \tag{14.172}$$

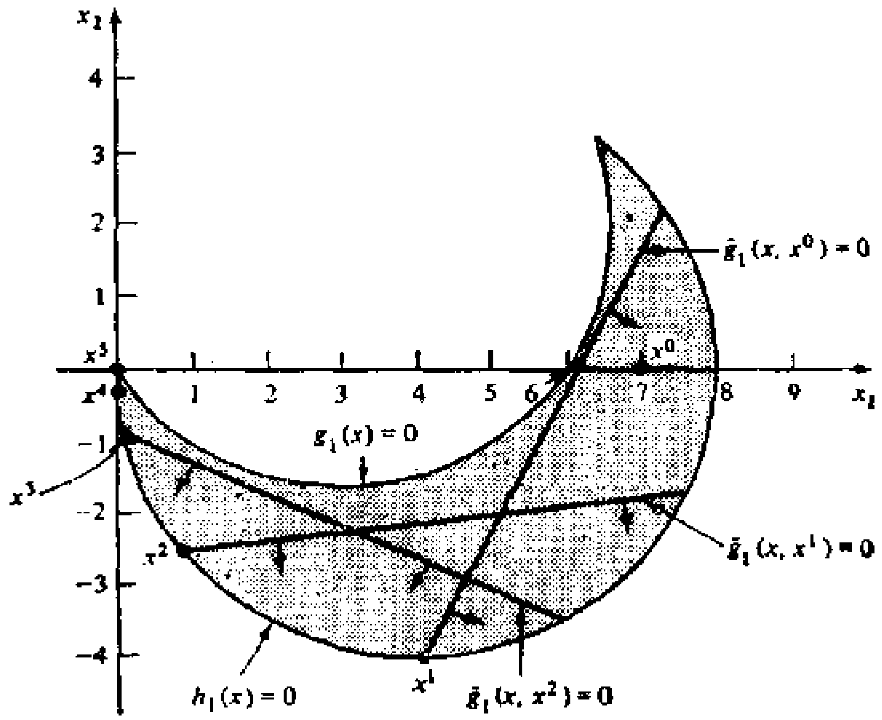


图 14.5 相补凸规划

受限制于约束

$$\tilde{g}_1(x, x^0)=8x_1-4x_2-49\geqslant 0, \tag{14.173}$$

$$h_1(x)=-(x_1-4)^2-(x_2)^2+16\geqslant 0. \tag{14.174}$$

由这个规划的解得到

$$A^0=\{(4.126, -3.998)\}. \tag{14.175}$$

因此，我们选取 $x^1=(4.126, -3.998)$，并在 x^1 展开 g_1，再求解所得到的凸规划等等.

表 14.1

迭　　代	x_1	x_2
0	7.000	0.000
1	4.126	−3.998
2	0.939	−2.575
3	0.094	−0.864
4	0.002	−0.134
5	0.000	−0.004

表 14.1 列出了这算法的迭代过程所产生的凸规划的最优解. 可以看到，在五次迭代中这些凸规划的最优解相当接近于一个局部极小值点.

初始点的选择，对于凸规划的解所收敛到的 K-T 点的位置有重要影响. 为了说明这个影响，假设同一例题的初始点取为 $x^0=(7.0, 2.1)$. 这时，通过三次迭代，这算法收敛于另一局部极小值点，如表 14.2 所示.

表 14.2

迭　　代	x_1	x_2
0	7.000	2.100
1	6 600	3.040
2	6.411	3.191
3	6.400	3.200

在这两种情形中，最优解集 A^* 由单个点组成. 为了说明最优解集可由多于一个点组成，选取 $x^0=(7, 2)$. 现在我们得到

$$\tilde{g}_1(x,\ x^0)=8x_1-53\geqslant 0, \tag{14.176}$$

且最优解集 A^0 为

$$A^0=\left\{(x_1,\ x_2): x_1=\frac{53}{8},\ |x_2|\leqslant\sqrt{\frac{15943}{8}}\right\}. \tag{14.177}$$

下一点 x^1 由 A^0 中任一点给出．如取

$$x^1=\left(\frac{53}{8},\ 2\right),$$

且求解 (CCP_1)，就得到一个新的集 A^1，它是一个线段，我们还可选取 $x_2^2=2$．在所有后面的迭代中总选取 $x_2^k=2$，最终收敛到第三个 K-T 点 $(3+\sqrt{13},\ 2)$，它不是一个局部极小值点． ∎

我们来检验相补凸规划算法的收敛性．首先注意，这个算法可以通过点到集合的映射 F 来描述，F 的半连续性在后面的定理中起重要作用．我们提醒读者，如果对每个 $x\in X\subset R^n$，可唯一确定 $Y\subset R^q$ 中的一个子集 $F(x)$，则 F 是一个点到集合的映射．换一种方式，如果 $\sigma(Y)$ 表示 Y 的子集之集合，那就可以说 F 把 X 映射到 $\sigma(Y)$ 中，采用记号：$F:X\to\sigma(Y)$．当然，X 和 Y 可以是同样的集合．与这种映射有关的一个重要性质是上半连续性[9]．点到集合的映射 $F:X\to\sigma(Y)$，称为在点 $\bar{x}\in X$ 是**上半连续的** (usc)，如果 $\{x^k\}\to\bar{x}$、$\{y^k\}\to\bar{y}$（即序列 $\{x^k\}$ 和 $\{y^k\}$ 分别收敛于 $\bar{x}$ 和 $\bar{y}$)、$x^k\in X$、$y^k\in F(x^k)\subset Y$ 就蕴涵着 $\bar{y}\in F(\bar{x})$．F 称为是上半连续的，如果它在每个点 $x\in X$ 是上半连续的．要求读者把这里定义的上半连续性，与第 6 章中引进的下半连续性作比较．

作为这个定义的一个直接结果，我们得出引理．

引理 14.2

设点到集合的映射 $F:X\to\sigma(Y)$ 在 $\bar{x}$ 是上半连续的．如果 $\{x^k\}\to\bar{x}$ 且 $y^k\in F(x^k)$，那么，$\{y^k\}$ 的每个凝聚点 y 必在集合 $F(\bar{x})$ 中．

【证明】 如果 y 是 $\{y^k\}$ 的一个凝聚点，那么存在一个子序列 $\{y^{k_i}\}$ 收敛到 y．因为 $y^{k_i}\in F(x^{k_i})$ 且序列 $\{x^{k_i}\}$ 是 $\{x^k\}$ 的一个子序列，因此 $\{x^{k_i}\}\to\bar{x}$．由 F 的上半连续性得到 $y\in F(\bar{x})$． ∎

其次，我们叙述并证明 Zangwill 定理[40]的稍作修改的形式.

定理 14.3

设点到集合的映射 $A: X \to \sigma(X)$ 确定了从 X 到 X 的子集族的一个映射，并设 X 是紧致集.

假设 A 满足下述条件:

(i) 如果 $x \in X$ 不是 (CCP) 的 K–T 点，那么 $y \in A(x)$ 蕴涵着 $f(y) < f(x)$. 如果 $x \in X$ 是 (CCP) 的 K–T 点，那么 $y \in A(x)$ 蕴涵着 $f(y) = f(x)$.

(ii) 映射 A 对于所有 $x \in X$ 是上半连续的.

那么，由 $x^0 \in X$, $x^{k+1} \in A(x^k)$ 产生的任何序列 $\{x^k\}$，必包含一个子序列收敛于 (CCP) 的 K–T 点.

【证明】 序列 $\{f(x^k)\}$ 是单调非增并有下界的(因为 X 是紧致集)，所以 $\lim\{f(x^k)\}$ 存在，把它记为 $\hat{f}$.

再由 X 的紧致性，序列 $\{x^k\}$ 必定包含一个收敛子序列，比如 $\{x^{k_i}\}$，具有极限 x^* 在 X 中. 对于这个子序列也有 $\lim\{f(x^{k_i})\} = \hat{f}$. 考察序列 $\{x^{k_i+1}\}$，因为它仍是 $\{x^k\}$ 的子序列，所以对于它还是有 $\lim\{f(x^{k_i+1})\} = \hat{f}$. 现在，$\{x^{k_i}\} \to x^*$、$x^{k_i+1} \in A(x^{k_i})$ 且 A 是上半连续的，因此由引理 14.2，$\{x^{k_i+1}\}$ 的每个凝聚点 x^{**} 在 $A(x^*)$ 中. 性质 (i) 将要求: 如果 x^* 不是 K–T 点，那么 $f(x^{**}) < f(x^*)$. 然而，因为 f 是连续的且 $\lim\{f(x^k)\} = \hat{f}$，所以我们有 $f(x^{**}) = \hat{f}$，并且也有 $f(x^*) = \hat{f}$. 因此点 x^* 必定是 K–T 点. 】

为了对于所提出的算法应用这个定理，我们必须指出这算法实际上是点到集合的映射，具有定理中所述的性质.

对于每个 $x^k \in X$，用线性约束 $\tilde{g}_i(x, x^k) \geqslant 0$ 来近似约束 $g_i(x) \geqslant 0$，可以抽象地看作一个映射 A_1，它把 x^k 变成系数集合

$$\{\nabla g_i(x^k),\ (x^k)^T \nabla g_i(x^k) - g_i(x^k)\}.$$

因为 g_i 是连续可微的，所以(单值)映射 A_1 是连续的，从而是上半连续的[9]. 凸规划 (CCP_k) 的最优解集合 $A(x^k)$，可以看作是系数集合上的一个函数. 抽象地，可以定义 A_2 为从系数集合到 R^n

中集合的一个点到集合的映射. 这样, $A=A_2A_1$ 是一个从 X 到 R^n 中子集的、复合的点到集合映射*. 我们将证明, 实际上它是由 X 到 X 的子集族的映射. 进一步我们还将证明, A 是上半连续的, 并且 $x^*\in X$ 是 (CCP) 的 K-T 点的充要条件为 $x^*\in A(x^*)$.

定理 14.4

映射 A 是从 X 到 $\sigma(X)$ 中的.

【证明】 对于 $x^k\in X$, 集合 $A(x^k)$ 是 (CCP_k) 的最优解集合. 对所有 $\hat{x}\in A(x^k)$, 有 $\hat{x}\in X(x^k)$. 即

$$g_i(x^k)+(\hat{x}-x^k)^T\nabla g_i(x^k)\geqslant 0, \quad i=1, \cdots, m, \qquad (14.178)$$

$$h_l(\hat{x})\geqslant 0, \quad l=1, \cdots, p. \qquad (14.179)$$

因为 g_i 是凸的, 所以

$$g_i(\hat{x})\geqslant g_i(x^k)+(\hat{x}-x^k)^T\nabla g_i(x^k), \qquad (14.180)$$

从而 $\hat{x}$ 满足 $g_i(\hat{x})\geqslant 0$. 因此 $\hat{x}\in X$, 即 $A(x^k)\subset X$. ■

下一个定理确立了定理 14.3 的性质 (i).

定理 14.5

向量 x^* 是 (CCP) 的一个 K-T 点的充要条件为 $x^*\in A(x^*)$.

【证明】 假设 x^* 是一个 K-T 点. 于是由定理 3.8, 存在乘子 λ^*、π^* 满足(3.72)至(3.74), 这些乘子的存在性, 等价于 x^* 是凸规划 (CCP_*) 之最优解的充要条件, (CCP_*) 是将(14.154)在 x^* 附近线性化而得到的, 也就是 $x^*\in A(x^*)$. 反之, 设 $x^*\in A(x^*)$, 则存在乘子满足(3.72)至(3.74). 这样, x^* 是(CCP) 的 K-T 点. ■

在建立了收敛性定理的性质 (i) 之后, 现在我们转到性质 (ii), 即验证 A 是上半连续的. 这里利用 Dantzig、Folkman、Shapiro[13] 和 Williams[35] 的下述结果, 已改写得适应我们的情况.

定理 14.6

假设对于任何 $\tilde{x}\in X$, 没有一个线性约束 $\tilde{g}_i(x, \tilde{x})\geqslant 0$ 是平凡的, 并且存在 $\hat{x}$ 使得

* 译注: 严格地说, $A=A_2A_1$ 是一个从 X 到 R^n 中子集族的复合的点到集映照. 原文的说法不够严格, 有若干处类似.

$$\tilde{g}_i(\hat{x}, \tilde{x}) > 0, \quad i=1, \cdots, m, \tag{14.181}$$

$$h_l(\hat{x}) > 0, \quad l=1, \cdots, p. \tag{14.182}$$

那么点到集合映射 $A_2(\tilde{x})$ 是上半连续的.

然后,我们叙述下述著名结果[9].

定理 14.7

如果 A_1 是 X 到 Y 中的上半连续映射, A_2 是 Y 到 Z 中的上半连续映射, 那么复合映射 $A=A_2A_1$ 是 X 到 Z 中的上半连续映射.

既然上述映射是上半连续的,那么,为了对复合映射 $A=A_2A_1$ 验证收敛性定理的性质 (ii),只需要证明,对于映射 A_2, 定理 14.6 的假设条件成立。

定理 14.8

线性约束 $\tilde{g}_i(x, \tilde{x}) \geqslant 0$ 是非平凡的.

【证明】 由本节开始时对 (CCP) 假设的正则性条件 (3), 约束 $g_i(x) \geqslant 0$ 是非平凡的, 即存在向量 $w \in R^n$ 使得 $g_i(w) < 0$. 于是从 g_i 的凸性可得

$$\tilde{g}_i(w, \tilde{x}) \leqslant g_i(w) < 0, \quad i=1, \cdots, m, \tag{14.183}$$

因而 $\tilde{g}_i(x, \tilde{x}) \geqslant 0$ 是平凡的. 】

现在必须证明, 存在 $\hat{x}$ 满足(14.181)和(14.182). 定理 4.19 在这里将证实是有用的.

定理 14.9

设 $\tilde{x}$ 在 X 中,则不存在 $\hat{x}$ 满足(14.181)和(14.182)的充要条件为: 存在数 $r_1, \cdots, r_m, q_1, \cdots, q_p$, 满足

$$r_i \geqslant 0, \quad i=1, \cdots, m \quad 和 \quad q_l \geqslant 0, \quad l=1, \cdots, p, \tag{14.184}$$

$$\sum_{i \in I(\tilde{x})} r_i + \sum_{l \in L(\tilde{x})} q_l > 0. \tag{14.185}$$

使得

$$\sum_{i \in I(\tilde{x})} r_i \frac{\partial g_i(\tilde{x})}{\partial x_j} + \sum_{l \in L(\tilde{x})} q_l \frac{\partial h_l(\tilde{x})}{\partial x_j} = 0, \quad j=1, \cdots, n \tag{14.186}$$

$$r_i g_i(\tilde{x}) = 0, \quad i=1, \cdots, m, \tag{14.187}$$

$$q_l h_l(\tilde{x}) = 0, \quad l=1, \cdots, p. \tag{14.188}$$

【证明】 假设不存在$\hat{x}$具有所要求的性质. 于是由定理4.19, 存在非负向量r和q, 使得对所有$x \in R^n$有

$$\sum_{i=1}^{m} r_i + \sum_{l=1}^{p} q_l > 0, \tag{14.189}$$

$$\sum_{i=1}^{m} r_i \left[g_i(\tilde{x}) + \sum_{j=1}^{n} \frac{\partial g_i(\tilde{x})}{\partial x_j}(x_j - \tilde{x}_j) \right] + \sum_{l=1}^{p} q_l h_l(x) \leqslant 0. \tag{14.190}$$

取$x=\tilde{x}$, 代入(14.190), 整理后得到

$$\sum_{i \in I(\tilde{x})} r_i g_i(\tilde{x}) + \sum_{l \in L(\tilde{x})} q_l h_l(\tilde{x}) + \sum_{i \bar{\in} I(\tilde{x})} r_i g_i(\tilde{x}) + \sum_{l \bar{\in} L(\tilde{x})} q_l h_l(\tilde{x}) \leqslant 0. \tag{14.191}$$

(14.191) 中前两项显然为零. 对于$i \bar{\in} I(\tilde{x})$和$l \bar{\in} L(\tilde{x})$, 分别有$g_i(\tilde{x}) > 0$和$h_l(\tilde{x}) > 0$. 这样, 对于$i \bar{\in} I(\tilde{x})$和$l \bar{\in} L(\tilde{x})$有$r_i = 0$和$q_l = 0$. 因此, (14.187)和(14.188)满足, 并对所有$x \in R^n$, (14.190)化简为

$$\sum_{j=1}^{n} \left[\sum_{i \in I(\tilde{x})} r_i \frac{\partial g_i(\tilde{x})}{\partial x_j} \right] (x_j - \tilde{x}_j) + \sum_{l \in L(\tilde{x})} q_l h_l(x) \leqslant 0. \tag{14.192}$$

显然, 不等式左端的函数在$\tilde{x}$达到它的无约束极大值, 因此它的一阶偏导数在$\tilde{x}$为零:

$$\sum_{i \in I(\tilde{x})} r_i \frac{\partial g_i(\tilde{x})}{\partial x_j} + \sum_{l \in L(\tilde{x})} q_l \frac{\partial h_l(\tilde{x})}{\partial x_j} = 0, \quad j=1, \cdots, n. \tag{14.193}$$

于是(14.186)成立.

现在假设存在向量$p \geqslant 0$、$q \geqslant 0$, 使得(14.184)至(14.188)成立, 并假设有一点$\hat{x}$满足(14.181)和(14.182). 因为h_l是凹的, (14.182)可由下式代替:

$$h_l(\tilde{x}) + \sum_{j=1}^{n} \frac{\partial h_l(\tilde{x})}{\partial x_j}(\hat{x}_j - \tilde{x}_j) > 0, \quad l=1, \cdots, p, \tag{14.194}$$

分别用相应的r_i和q_l乘(14.181)和(14.194)中的每个不等式, 得到

$$\sum_{j=1}^{n}\left[\sum_{i=1}^{m} r_i \frac{\partial g_i(\tilde{x})}{\partial x_j}+\sum_{l=1}^{p} \frac{\partial h_l(\tilde{x})}{\partial x_j}\right](\hat{x}_j-\tilde{x}_j)$$
$$+\sum_{i=1}^{m} r_i g_i(\tilde{x})+\sum_{l=1}^{p} q_l h_l(\tilde{x})>0, \qquad (14.195)$$

这与假设矛盾. ∎

比较定理 14.9 的条件和本节开始时对 (CCP) 假设的正则性条件(2)，我们断定存在一个点 $\hat{x}$ 具有所要求的性质. 这样，定理 14.7 至定理 14.9 建立了映射 A 的上半连续性, 从而定理 14.3 的性质 (ii) 是满足的. 这就完全证明了为求解相补凸规划所提出的算法的收敛性.

显然, 作一些容易的修改后, 这个方法可以推广到具有广义凸函数和广义凹函数的规划. 另一推广是通过对约束(14.154)构造支撑超平面发展起来的, 这类似于上一节的 Veinott 方法. 当然，那时 g_i 可以是拟凸的而不一定是凸的.

也许, 刚才所述的相补凸规划算法的最重要应用, 在于求解第 7 章曾提到的正项式规划或相补几何规划问题.

考虑问题

$$\text{(CGP)} \qquad \min x_0 \qquad (14.196)$$

受限制于

$$P_i(x)-Q_i(x)\leqslant 1, \quad i=1, \cdots, m, \qquad (14.197)$$

其中 $x=(x_0, x_1, \cdots, x_n)^T$, P_i、Q_i 是正项式. 可以证明，这种规划能变换为“相补 (log, log)-凸”规划(参见第 6、7 章)，其中 (log, log)-凸函数和 (log, log)-凹函数出现在约束中. 当然, 这种规划也可以变换为一个等价的相补凸规划.

关于概念性方法的细节，读者可参考 Avriel[2] 和 Avriel、Williams[5,6]. Dembo[14] 已成功地实现了这个方法，他用割平面算法去解凸的子问题或 (log, log)-凸的子问题. 完整的算法叙述在 Avriel、Dembo、Passy[3] 中. 在练习 14.M 中将看到，用相补凸规划方法还能处理更一般的问题, 它们包含正项式的某种函数，而不是(14.197)中正项式之差.

练　　习

14.A. 继续做例题 14.1.1. 用 Griffith-Stewart 方法进行两次以上的迭代，并画出子问题的约束. 解析地计算原来问题的最优解，并观察收敛于这个解的情况。

14.B. 画出 Fletcher 的超方体法的框图，并对下面的问题至少进行两次迭代，以说明这个方法：

$$\min f(x)=\left(x_1+\frac{1}{x_2}\right)^{1/2}+\frac{4}{(x_1-x_2)^2} \tag{14.198}$$

受限制于

$$x_1+x_2\geqslant 2, \tag{14.199}$$

$$x_1\geqslant 0,\quad x_2\geqslant 0, \tag{14.200}$$

按照算法的需要自己选择参数.

14.C. 现在来看一下如何对非线性规划算法构造试验问题，其最优解是预先已知的. 我们将利用 Rosen、Suzuki[30] 提出的构造法. 假设要求解问题

$$\min f(x)=\phi(x)+c^Tx,\quad x\in R^n \tag{14.201}$$

受限制于

$$g_i(x)=\psi_i(x)+b_i\geqslant 0,\quad i=1,\cdots,m, \tag{14.202}$$

其中 ϕ 和 $\psi_1,\cdots,\psi_m$ 分别是给定的可微实严格凸函数和凹函数. 选取任意 $x^*\in R^n$ 为预先决定的最优解，并选取 $\lambda^*\in R^m,\lambda^*\geqslant 0$ 作为相应的 Kuhn-Tucker 乘子，选取 $b_i,\ i=1,\cdots,m$, 使得

$$g_i(x^*)=0,\quad 若\ \lambda_i^*>0, \tag{14.203}$$

$$g_i(x^*)\geqslant 0,\quad 若\ \lambda_i^*=0. \tag{14.204}$$

用下式计算向量 c:

$$c=\sum_{i=1}^{m}\lambda_i^*\nabla\psi_i(x^*)-\nabla\phi(x^*), \tag{14.205}$$

证明参数 b 和 c 的这种选择确实保证了 x^* 是所希望的最优解. 指出 ϕ 的凸性和 ψ 的凹性怎样才可以用另外的条件代替，使得这个构造方法仍然可用.

14.D. 用上题的方法构造一个具有线性约束的非二次试验问题，对它应用既约梯度法进行几次迭代.

14.E. 验证(14.89)和(14.90)是问题 (NEP) 的最优性 Kuhn-Tucker 必要

条件.

14.F. 画出 GRG 方法的框图. 用练习 14.C 的 Rosen-Suzuki 方法构造一个两变量非线性约束试验问题, 应用 GRG 方法解这问题, 并画出最优化的行程.

14.G. 假设我们求单变量非线性方程 $f(x)=0$ 的解, 其中 f 是一个严格递增的可微凹函数. 把这个问题转化为适于用割平面法求解的非线性规划问题. 试找出割平面法所产生的点与 Newton 法直接用于方程求解问题时所生成的点之间的关系.

14.H. 画出例 14.3.1 中非线性规划的图, 用视察法求最优解. 画出增加的割平面, 并注意算法的进程.

14.I. 应用支撑超平面法于例 14.3.1 的问题. 至少进行两次增加平面的步骤.

14.J. 对于 (h, ϕ)-凸规划导出"割曲面"方法. 用 (log, log)-凸的几何规划问题求解来说明你的方法.

14.K. 第 14.4 节的相补凸规划方法, 可以看作是"割平面"方法(为什么?). 把正文中提出的方法修改为一种"支撑超平面"方法. 对例题 14.4.1 中的问题试一下你的方法.

14.L. 对规划中的函数和变量作必要的变换, 并在需要时增加附加变量和约束, 从而把相补几何规划叙述为相补凸规划. 先把例 14.4.1 中的问题写成相补几何规划, 然后用你的推导再把它写成相补凸规划. 注意, 近似规划是原来变量的普通几何(正项式)规划.

14.M. 给定非线性规划

$$\text{(AP)} \qquad \min \sum_{i\in I_0} w_{i0}\left|\frac{P_{i0}(x)-Q_{i0}(x)}{R_{i0}(x)}\right|^{a_{i0}} \tag{14.206}$$

受限制于

$$\sum_{i\in I_k} w_{ik}\left|\frac{P_{ik}(x)-Q_{ik}(x)}{R_{ik}(x)}\right|^{a_{ik}} \leqslant 1, \quad k=1, \cdots, m, \tag{14.207}$$

其中 P_{ik}、Q_{ik}、R_{ik} 是正项式, 且对于 $i\in I_k$, $k=0, \cdots, m$, w_{ik}、a_{ik} 是给定的正数. 证明用增加附加变量和约束, (AP) 可以转化为等价的相补几何规划. 叙述近似的几何规划并尝试消去增加的附加变量.

参 考 文 献

1. ABADIE, J., and J. CARPENTIER, "Generalization of the Wolfe Reduced Gradient Method to the Case of Nonlinear Constraints," in *Optimization*, R. Fletcher (Ed.), Academic Press, London, 1969.
2. AVRIEL, M., "Methods for Solving Signomial and Reverse Convex Programming Problems," in *Optimization and Design*, M. Avriel, M. J. Rijckaert, and D. J. Wilde (Eds.), Prentice-Hall, Englewood Cliffs, N.J., 1973.
3. AVRIEL, M., R. S. DEMBO, and U. PASSY, "Solution of Generalized Geometric Programs," *Int. J. Numer. Methods in Engineering*, 9, 149–168 (1975).
4. AVRIEL, M., and A. C. WILLIAMS, "Complementary Convex Programming," Mobil R & D Corp. Central Research Div. Progress Memorandum, Princeton, N.J., May 1968.
5. AVRIEL, M., and A. C. WILLIAMS, "Complementary Geometric Programming," *SIAM J. Appl. Math.*, 19, 125–141 (1970).
6. AVRIEL, M., and A. C. WILLIAMS, "An Extension of Geometric Programming with Applications in Engineering Optimization," *J. Eng. Math.*, 5, 187–194 (1971).
7. BEALE, E. M. L., "Numerical Methods," in *Nonlinear Programming*, J. Abadie (Ed.), North-Holland Publishing Co., Amsterdam, 1967.
8. BEALE, E. M. L., "Nonlinear Optimization by Simplex-like Methods," in *Optimization*, R. Fletcher (Ed.), Academic Press, London, 1969.
9. BERGE, C., *Topological Spaces*, Oliver & Boyd Ltd., Edinburgh, 1963.
10. CHENEY, E. W., and A. A. GOLDSTEIN, "Newton's Method for Convex Programming and Tchebycheff Approximation," *Numer. Math.*, 1, 253–268 (1959).
11. COLVILLE, A. R., "A Comparative Study on Nonlinear Programming Codes," *Proceedings of the Princeton Symposium on Mathematical Programming*, H. W. Kuhn (Ed.), Princeton University Press, Princeton, N.J., 1970.
12. DANTZIG, G. B., *Linear Programming and Extensions*, Princeton University Press, Princeton, N.J., 1963.
13. DANTZIG, G. B., J. FOLKMAN, and N. Z. SHAPIRO, "On the Continuity of the Minimum Set of a Continuous Function," *J. Math. Anal. & Appl.*, 17, 519–548 (1967).
14. DEMBO, R. S., "Solution of Complementary Geometric Programming Problems," M.Sc. thesis, Technion, Israel Institute of Technology, Haifa, 1972.
15. EAVES, B. C., and W. I. ZANGWILL, "Generalized Cutting Plane Algorithms," *SIAM J. Control*, 9, 529–542 (1971).
16. FLETCHER, R., "A General Quadratic Programming Algorithm, *J. Inst. Math. Appl.*, 7, 76–91 (1971).

17. Fletcher, R., "An Algorithm for Solving Linearly Constrained Optimization Problems," *Math. Prog.*, **2**, 133–165 (1972).

18. Goldfeld, S. M., R. E. Quandt, and H. F. Trotter, "Maximization by Quadratic Hill-Climbing," *Econometrica*, **34**, 541–551 (1966).

19. Griffith, R. E., and R. A. Stewart, "A Nonlinear Programming Technique for the Optimization of Continuous Processing Systems," *Management Science*, **7**, 379–392 (1961).

20. Kelley, J. E., "The Cutting-Plane Method for Solving Convex Programs," *J. SIAM*, **8**, 703–712 (1960).

21. Levenberg, K. A., "A Method for the Solution of Certain Nonlinear Problems in Least Squares," *Quart. Appl. Math.*, **2**, 164–168 (1944).

22. Levitin, E. S., and B. T. Polyak, "Constrained Minimization Methods," *USSR Comp. Math. & Math. Phys.*, **6**, **5**, 1–50 (1966).

23. Marquardt, D. W., "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *J. SIAM*, **11**, 431–441 (1963).

24. McCormick, G. P., "Anti-zig-zagging by Bending," *Management Science*, **15**, 315–320 (1969).

25. McCormick, G. P., "The Variable Reduction Method for Nonlinear Programming," *Management Science*, **17**, 146–160 (1970).

26. Pearson, J. D., "Variable Metric Methods of Minimisation," *Computer J.*, **12**, 171–178 (1969).

27. Powell, M. J. D., "A New Algorithm for Unconstrained Minimization," in *Nonlinear Programming*, J. B. Rosen, O. L. Mangasarian, and K. Ritter (Eds.), Academic Press, New York, 1970.

28. Powell, M. J. D., "Convergence Properties of a Class of Minimization Algorithms," Report C.S.S. 8, A.E.R.E. Harwell, United Kingdom, April 1974.

29. Rosen, J. B., "Iterative Solution of Nonlinear Optimal Control Problems," *SIAM J. Control*, **4**, 223–244 (1966).

30. Rosen, J. B., and S. Suzuki, "Construction of Nonlinear Programming Test Problems," *Comm. of the ACM*, **8**, 113 (1965).

31. Topkis, D. M., "Cutting-Plane Methods without Nested Constraint Sets," *Operations Research*, **18**, 404–413 (1970).

32. Veinott, A. F., "The Supporting Hyperplane Method for Unimodal Programming," *Operations Research*, **15**, 147–152 (1967).

33. Wilde, D. J., "Jacobians in Constrained Nonlinear Optimization," *Operations Research*, **13**, 848–856 (1965).

34. Wilde, D. J., and C. S. Beightler, *Foundations of Optimization*, Prentice-Hall, Englewood Cliffs, N.J., 1967.

35. Williams, A. C., "Marginal Values in Linear Programming," *J. SIAM*, **11**, 82–94 (1963).

36. WILSON, R. B., "A Simplicial Algorithm for Concave Programming," Doctoral dissertation, Harvard University, Cambridge, Mass., 1963.

37. WOLFE, P., "Methods of Nonlinear Programming," in *Recent Advances in Mathematical Programming*, R. L. Graves and P. Wolfe (Eds.), McGraw-Hill Book Co., New York, 1963.

38. WOLFE, P., "Methods of Nonlinear Programming," in *Nonlinear Programming*, J. Abadie (Ed.), North-Holland Publishing Co., Amsterdam, 1967.

39. WOLFE, P., "Convergence Theory in Nonlinear Programming," in *Integer and Nonlinear Programming*, J. Abadie (Ed.), North-Holland Publishing Co., Amsterdam, 1970.

40. ZANGWILL, W. I., "Convergence Conditions for Nonlinear Programming Algorithms," *Management Science*, 16, 1–13 (1969).

41. ZANGWILL, W. I., *Nonlinear Programming: A Unified Approach*, Prentice-Hall, Englewood Cliffs, N.J., 1969.

索　　引

（"上"表示上册页码，"下"表示下册页码）

三　划

四　划

五 划

七　划

八　划

九　划

十 划

十 一 划

十 二 划

十 三 划

十四——十六划